THE ROLE OF GLYOXYLIC ACID IN THE CHEMISTRY OF THE ORIGIN OF LIFE

A Dissertation
Presented to
The Academic Faculty

by

Christopher J. Butch

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in Chemical Engineering in the
School of Chemical and Biomolecular Engineering

Georgia Institute of Technology
December, 2014

THE ROLE OF GLYOXYLIC ACID IN THE CHEMISTRY OF THE ORIGIN OF LIFE

Approved by:

Dr. Charles Eckert, Advisor
School of Chemical and Biomolecular Engineering
Georgia Institute of Technology

Dr. David Sholl
School of Chemical and Biomolecular Engineering
Georgia Institute of Technology

Dr. Charles Liotta, Advisor
School of Chemical and Biomolecular Engineering/ School of Chemistry
Georgia Institute of Technology

Dr. Thomas Orlando
School of Chemistry
Georgia Institute of Technology

Dr. Ramanayaranan Krishnamurthy
Department of Chemistry
The Scripps Research Institute

Dr. Facundo Fernandez
School of Chemistry
Georgia Institute of Technology

Date Approved: 8/19/2014

To my wife, Diane. Without your love and support, this may have been a TPS report.
Thank you.

ACKNOWLEDGEMENTS

I would like to acknowledge the three excellent undergraduates who I have had the privilege to work with and mentor. The diligent efforts of Jacob Crowe, Rebeca Vindas, and Sara Mills were crucial to the work presented herein. Pamela Pollet deserves hearty thanks for her role as lab manager and research mentor. I would like to thank Deborah Babykin, Michele Yager, and Shantel Floyd for their roles in administrative support of our group; I am certain without these three women the lab and possibly the school would grind immediately to a halt. Leslie Gelbaum and Facundo Fernandez were invaluable mentors in helping me to learn the analytical techniques of NMR and mass spectrometry, respectively. Elizabeth Cope was responsible for acclimating me both to the laboratory, and the origin of life field. Finally, all my colleagues from the Eckert-Liotta group and the Center for Chemical Evolution have provided invaluable support, advice and encouragement. Finally, I would like to thank the Center for Chemical Evolution, NSF, and NASA, for the funding of the work presented. Thank you all.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

DAMN	Diaminomaleonitrile
DFT	Density Functional Theory
DHA	Dihydroxyacetone
DHBA	Dihydroxybutanoic Acid
DHF	Dihydroxyfumarate
DKP	Diketopiperazine
DLH	Double Layered Hydroxide
ESI-MS	Electrospray Injection Mass Spectrometry
FDCA	2,5-Furandicarboxylic acid
HCHO	Formaldehyde
HCN	Hydrogen Cyanide
HMF	Hydroxymethylfurfuraldehyde
HMTA	Hexamethylenetetramine
kcal	Kilocalorie
LA	Levulinic Acid
LCMS	Liquid Chromatography Mass Spectrometry
mol	Mole
mmol	Millimole
MRM	Multiple Reaction Monitoring
NHC	N-Heterocyclic Carbene
NPH	Nitrophenylhydrazine
PCM	Polarizable Continuum Model
RNA	Ribonucleic Acid
$scCO_2$	Super Critical Carbon Dioxide

SUMMARY

Herein, I present detailed mechanistic analysis on the chemistry of glyoxylate as it pertains to forming biologically relevant molecules on the Hadean Earth. Chemistry covered includes: 1) the dimerization of glyoxylate to form dihydroxyfumarate(DHF), a heretofore unknown reaction, important to substantiating Eschenmoser's glyoxylate scenario. 2) Formation of sugars from polymerization of glyoxylate. 3) Formation of tartrate and sugar acids from high pH reactions of DHF. 4) Formation of glycine polypeptides from glyoxylate by transamination and coupling promoted by hexamethylenetetramine. 5) Formation of glyoxylate under conditions which could be plausibly found on the early earth.

CHAPTER 1:
INTRODUCTION TO PREBIOTIC CHEMISTRY

The question “How did life come to be?” is simultaneously one of the most interesting and most intractable unanswered questions in science. The complexity of life begs for an explanation other than simple chance. Even viroids, which could be considered unimolecular life, are incredibly intricate[1]. The smallest of these weighs in at just 220 nucleosides[2], some eight thousand atoms, arranged just so. These 220 bases combine to form an RNA chain which has a spherical, protective shape; catalytic activity; and all the information to co-opt a host cell. It is a byzantine piece of molecular machinery, and even so, it is still dependant on the workings of a host cell to reproduce. To obtain independent reproduction, the most defining characteristic of life, requires a thousand-fold increase in genetic material[3], and all the accompanying accoutrements of celldom; machinery to produce membranes, to produce proteins, and to reproduce. There is simply no way that all of these pieces of molecular machinery could spontaneously spring into existence, assemble, and come together to form life, and this is, of course, because they didn’t.

It is very unlikely that the origin of life can be distilled to one discrete event wherein previously inert RNA, proteins, and lipids suddenly and spontaneously coalesced into a working bacterium. Instead, the very first “living organism” was probably similar to the viroid described above; a single piece of RNA (or similar informational polymer) that through environmental circumstance was able to duplicate, creating a population that could then evolve, eventually achieving beneficial function. Once this first functional, replicating polymer existed, Darwinian evolution can take hold, selecting for ever greater fitness[4]. Ultimately, these functioning RNAs could be concentrated through any of numerous geochemical processes, and encapsulated in fatty acids, creating the first cells[5]. These newly formed cells would create discrete populations of RNA, drastically

increasing the pace of evolution[6]. From this time on, these encapsulated populations would spread and adapt to different chemistries, evolving to take advantage of local resources, and ultimately, with the evolution of the ribosome, begin to control their own environments through the production of proteins[7]. With this assembly of proteins, RNA, and membrane, all the necessary strata are assembled, upon which, given a few billion years, all the mechanisms of modern life can be constructed. This progression of RNA organisms to modern life is the theory of the "RNA World," and is the origin story which resonates most strongly within the origin of life community[8]. While there are numerous unknowns related to the specifics of these processes, the RNA world theory presents a coherent and plausible method for life to arise from the earliest informational polymers.

While the RNA world theory addresses the origin of life after the advent of functional polymers, theories for what happened prior to the RNA world are less defined. At the beginning of this world prior to RNA, the chemical inventory on the young cooling Earth would have been extremely simple, likely mirroring the chemical composition of meteorites which may have served as early delivery systems[9]. From this dilute early ocean inventory, some process must have occurred to transform simple available molecules such as CO_2 and ammonia, into the carbohydrates and nucleobases necessary to form an RNA-like polymer. It is within this context of this time preceeding the RNA world that this dissertation will discuss the viability of glyoxylic acid as a prebiotic reactant to produce a number of biogenic molecules.

The prebiotic chemistry prior to the RNA world took place over millions of years, and there is a large volume of literature cataloging the numerous reactions that may have taken place leading to the RNA world. Within this literature the chemistry of formaldehyde and the chemistry of cyanide are invoked repeatedly and an understanding

of this chemistry is essential both to understanding prebiotic chemistry in general, and to understand the chemical advantages offered by glyoxylate in the prebiotic context.

1.1 Prebiotic Chemistry of Formaldehyde

The immense interest in the prebiotic chemistry of formaldehyde is driven by a number of factors. Formaldehyde (HCHO) is the simplest aldehyde, as well as the second reduction product of carbon dioxide and second oxidation product of methane. As such is expected to have been produced on early earth regardless of atmospheric composition [10-12]. In fact, Pinto estimated that as a result of ready atmospheric formation, the steady state concentration of formaldehyde in the prebiotic ocean could have been as high as 1mM[13].Formaldehyde has also been observed as a major component of the interstellar matrix[14-18] and has been shown in the laboratory to form under simulated interstellar conditions[19, 20]. Consequently, comets, meteorites, and interstellar grains are also thought to be a potential supply vector for Archean formaldehyde [21-23]. In addition to these atmospheric and exogenic means of formaldehyde productions, aqueous chemical pathways also exist. Formaldehyde has been shown to be produced in small amounts by aqueous photo-reduction of CO_2 in the presence of Fe^{2+}[24] and is a known intermediate in the production of methane by serpentinization[25] with low concentrations observed in modern hydrothermal systems[26].

Setting aside its presumed ready availability, what likely generates the most interest in formaldehyde in the prebiotic community is its role in the Butlerow synthesis of carbohydrates, a.k.a. the formose reaction[27]. In 1861, Butlerow observed formation of sugar species when he subjected solutions of formaldehyde to mildly alkaline conditions, however it was nearly a century before a reasonable mechanism was proposed

for this reaction. In 1959, Breslow proposed an autocatalytic series of aldol and retro aldol reactions to be the primary drivers of the formose reaction[28]. In this mechanism it is glycolaldehyde, rather than formaldehyde which is driving sugar production. Breslow did not suggest a mechanism for this initial dimerization, simply calling the required time an incubation period. It has since been shown that this incubation period is reflective of the amount of initial carbohydrate impurities, likely glycolaldehyde; the homogeneous formose reaction does not proceed when conducted with highly purified formaldehyde[29], and the reaction kinetics are proportional to the initial glycolaldehyde (or other carbohydrate) concentration[30]. Presence of an initial carbohydrate is not completely required, however, as there is evidence that the dimerization of formaldehyde can be photocatalyzed[31] or occur on the surface of specific minerals[32]. Regardless, the necessity of glycolaldehyde for the reaction to progress is strong evidence toward Breslow's mechanism being correct, a fact which has been further substantiated by numerous works[33-35]

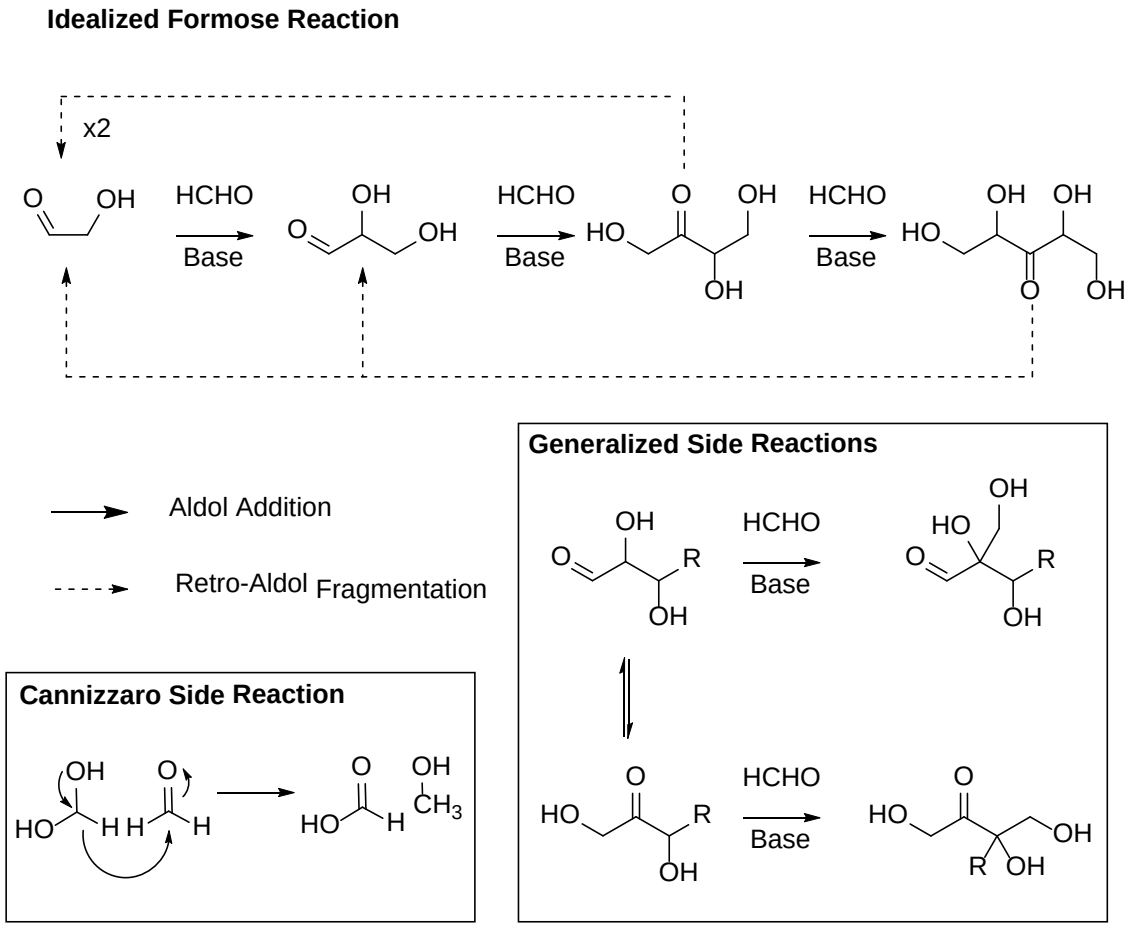


Figure 1 – Mechanism of the Formose Reaction In the ideal reaction formaldehyde additions occur only at terminal positions resulting in the production of linear ketoses. Side reactions including carbonyl migrations and branched additions greatly complicate the reaction mixture. An additional side reaction, the disproportionation of two formaldehyde to a formate and methanol (the Cannizzaro reaction) is also shown. This reaction can be responsible for consuming a significant amount of available formaldehyde, depending on reaction conditions. Not shown are aldol reactions involving species other than formaldehyde, which are an additional complicating factor.

A viable route to linear sugars is crucial to the RNA world hypothesis, due to the important role of phosphophorylated ribose as the structural backbone of modern genetic material[36].The “state of the art” in the synthesis of ribose from formaldehyde is a 22% yield of pentuloses [37] which could potentially isomerize to ribose, or be used as an alternate backbone material[38]. To achieve these yields however, three separate mineral conditions are invoked. First calcium hydroxide (or other divalent hydroxide) must be present as a catalyst, a well-known dependency of the formose reaction[39-41] . Second,

borate minerals must be present to guide the reaction through stabilization of tetroses and pentoses [42]. Third molybdates must be present to promote rearrangement of branched products to linear sugars[43]. While this reaction represents tremendous progress in demonstrating prebiotic carbohydrate synthesis in support of the RNA world, the mineral complexity required keeps the door open for a simpler chemical scenario.

In addition to this potential role in carbohydrate formation, formaldehyde can react with ammonia and formaldehyde in the Strecker mechanism (Figure 2), as a potential source of prebiotic glycine and serine (from glycolaldehyde)[44]. Both ammonia [45, 46] and cyanide[47, 48] are expected to have been present in relatively large amounts on the early earth, making Strecker synthesis a likely prebiotic process[49]. While functional proteins are not thought to have played an important role in the RNA world hypothesis until after the evolution of the ribosome, peptides must have been present at this time for the selection to even occur [50]. Further, glycine is known to catalyze both the formation of carbohydrates[51] and polypeptides[52]. Glycine has also been shown to react in formamide to yield adenine and purine [53], while serine has shown as a backbone in the formation of peptide nucleic acids, a potential early alternative informational polymer[54, 55] as have modified glycines[56, 57].

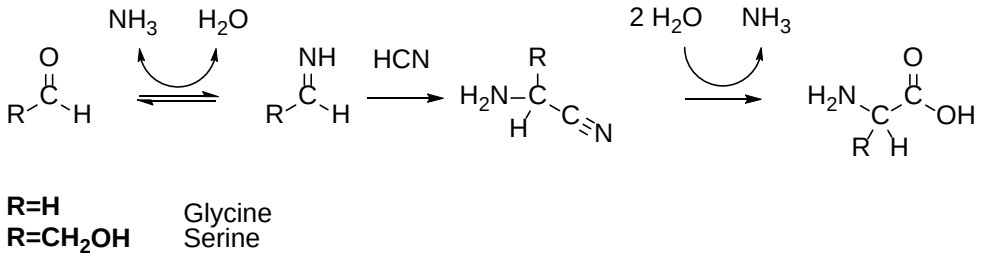


Figure 2 – Strecker Synthesis of Glycine and Serine from Formaldehyde and Glycolaldehyde In this mechanism the aldehyde reacts initially with ammonia to form the analogous imine species. This imine is attacked by cyanide to form an amino nitrile intermediate. Hydrolysis of the nitrile yields the amino acid.

While this is by no means a comprehensive accounting of the prebiotic chemistry of formaldehyde, it does demonstrate that formaldehyde has a role in the production of each of the three major biological monomers: amino acids, carbohydrates, and nucleobases. It is this chemical versatility combined with ready formation that defines the prebiotic plausibility of formaldehyde as a quintessential prebiotic source molecule.

1.2 Prebiotic Chemistry of Cyanide

Hydrogen cyanide is the simplest nitrile, and like formaldehyde, is a molecule expected to be prebiotically ubiquitous. This expectation is partially due to numerous interstellar observations of cyanide[58-60] as well as in the atmosphere of Jupiter's moon, Titan,[61-63] which is often used as a model for prebiotic earth[64-66]. Chameides and Walker calculated annual production of HCN by electrical discharge in a number of possible prebiotic atmospheres and showed likely production in all of them, albeit at a low rate in CO_2[67]. This work was experimentally substantiated by Stribling and Miller[47] showing a yield significantly higher than predicted in the CO_2 atmosphere. HCN can also be produced photochemically by generation of nitrogen atoms in the upper atmosphere and subsequent reaction with methane[68].

Delivery of HCN to earth on meteors is also a viable means of supply. Based on the inventory of ammonia containing compounds in the Murchison meteorite and known kinetics of formation of these compounds it has been estimated that the concentration of HCN on the meteorite could have been as high as 5mM (37-75ug/gm)[69]. The estimated flux of carbonaceous chondrites, the type of meteor including the Murchison meteorite, during the late heavy bombardment is $8x10^7$ kg/yr yielding an upper estimate of 6000 kg/yr of cyanide or .000045nmol cm^{-2} yr^{-1} . If one assumes interplanetary dust particles

(flux of $3x10^{12}$ kg/yr) have the same composition as carbonaceous condrites, as is often done[70], this estimated cyanide flux rises to 1.69 nmol cm^{-2} yr^{-1}.

Once produced, cyanide can undergo numerous chemistries of prebiotic interest, the most pronounced of which is the tendency to polymerize. The first step of this polymerization is the addition of a cyanide ion to neutral cyanide to form iminoacetonitrile[71]. This iminoacetonitrile can also dimerize to form ultimately diaminomaleonitrile (DAMN)[72]. Alternately, a cyanide ion can react with iminoacetonitrile to form aminomalononitrile, the HCN trimer, which can react further with cyanide to form insoluble tars[73]. While these polymers are not of specific prebiotic interest, hydrolysis of the imino groups can directly yield a polymer with peptide linkages incorporated without the need for amino acid synthesis[74]. In addition to this direct synthesis of protein-like polymers, it has been shown that DAMN can promote formation of peptide bonds between amino acid monomers[75].

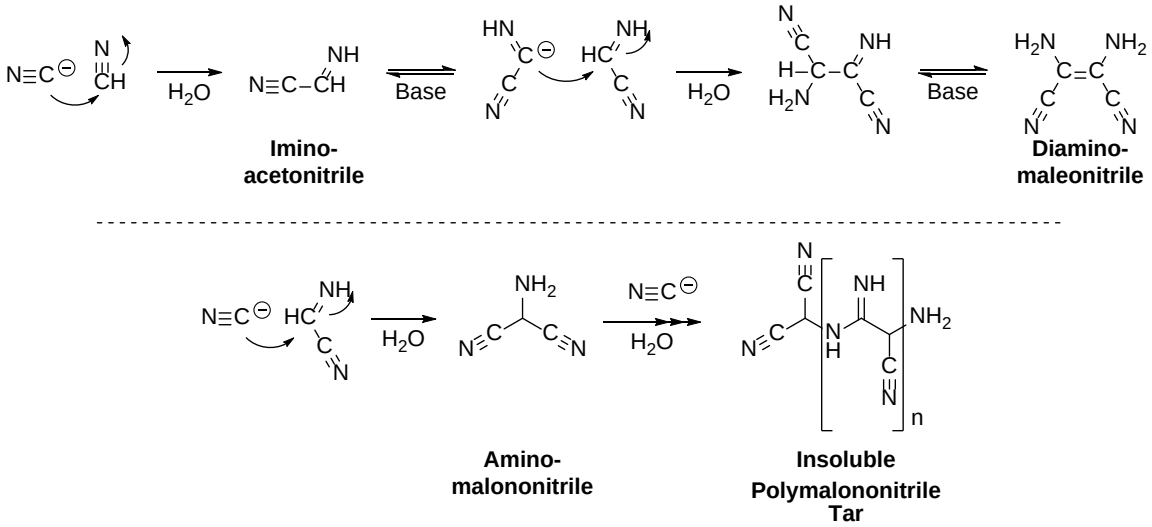


Figure 3 – Formation of the Dimer, Tetramer, and Polymer of Hydrogen Cyanide

In addition to these roles in the formation of polypeptides, reaction of cyanide is the most straightforward route to a number of nucleobases, as first shown by Oró[76]

with the production of adenine (Figure 4). In this reaction one equivalent of malononitrile and two equivalents of cyanide in the presence of ammonia are converted to their corresponding amidines. These amidines condense to form adenine, for a net reaction stoichiometry of five equivalents of cyanide to one adenine. Subsequent works have shown synthesis of hypoxanthine [77], uracil[78], and guanine[79] via similar pathways. Formamide can also originate from partial hydrolysis of cyanide[80], and can react to form several nucleic acids in its own right[81].

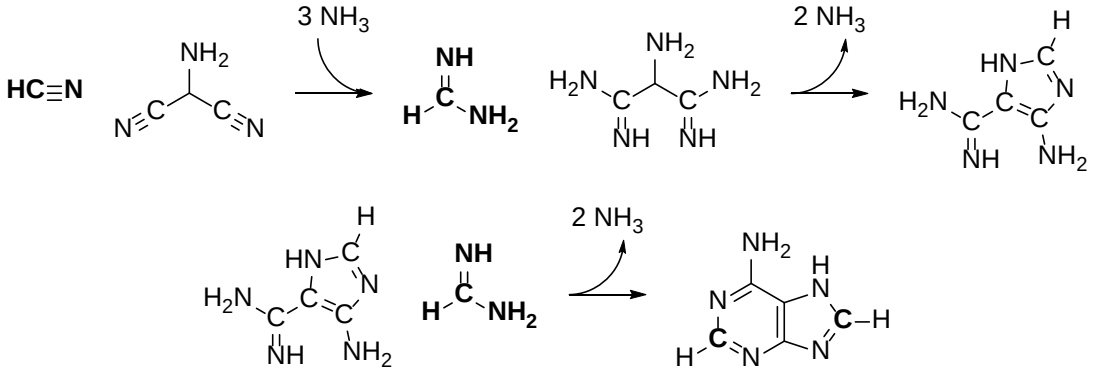


Figure 4 – Mechanism of Adenine Synthesis from Cyanide and Malononitrile by way of the Analogous Amidines. Two equivalents of cyanide and one malononitrile are converted to the corresponding amidines by reaction with ammonia. A series of condensation reactions results in adenine with the bridging carbons provided by formamidines shown in bold.

Cyanide and formaldehyde can also react with one another to form glycolonitrile, another prebiotically interesting species[82]. Presence of this species increases the yield of adenine from cyanide discussed above by a factor of 5. Sutherland proposes that photochemical reduction of glycolonitrile to glycolaldehyde could be a means of bypassing the troublesome initial dimerization of formaldehyde in the formose reaction[83]. The most notable feature of glycolonitrile is its tendency to polymerize under neutral and basic conditions to form amino-oxazoles and amino-oxazolines which

may have been among the first heterocycles[84]. When this polymerization occurs in the presence of excess cyanide 8-methyl-adenine is produced as an additional product [85].

1.3 The Glyoxylate Scenario

Glyoxylic acid and glyoxylate have seen limited investigations in the context of prebiotic chemistry dating back to the 1950s. In 1957, Braunshtein [86] suggested that glyoxylic acid could serve 1) as a catalyst in the formation of levulinic acid, in a method analogous to chemistry demonstrated by Metzler[87], and 2) substitutions of serine to form other amino acids. This suggestion that glyoxylate could play a tangential role in the synthesis of a citric acid cycle intermediates and amino acids and thus bolster the theory of prebiotic metabolism essentially defined the role of glyoxylic acid in the eyes of the prebiotic community for the next 50 years; the few papers that have investigated prebiotic glyoxylate chemistry have done so in terms of metabolic systems[88, 89] and amino acid synthesis[90, 91]. In recent years, however, the hypothesis that glyoxylate may be an important prebiotic building block for the prebiotic synthesis of biological molecules has been gaining ground in the prebiotic community.

In 2006, in the first demonstration of glyoxylate as a prebiotic chemical building block, Bean et al. synthesized glyoxylate-linked nucleic acids demonstrating that glyoxylate hemiacetals could replace phosphate esters in the RNA backbone[92]. This was followed by Eschenmoser's seminal paper on the "glyoxylate scenario"[93], wherein he theorized that glyoxylic acid may have played a crucial role in the development of a prebiotic metabolism. In this scenario glyoxylic acid is a primary building block, rather than catalyst or intermediate, in numerous pathways leading to biogenic molecules such as sugars, amino acids, and nucleobases. Of specific interest is the potential for the

"glyoxylate scenario" to yield linear sugars, crucial to modern genetic and cellular structure. The selective production of linear sugars in Eschenmoser's scenario may help explain the modern ubiquity of RNA and the tricarboxylic acid cycle, both of which depend on linear sugars, by providing a ready source of these sugars on hadean Earth. In his paper, Eschenmoser proposed that glyoxylate could dimerize to form dihydroxyfumarate which can then undergo further reaction with small aldehydes and ketones which could then decarboxylate yielding linear sugars. In 2011, Sagi et al. showed that DHF self-condenses into pentulosonic acid, a sugar acid species that can decarboxylate to form erythrulose, thus laying an experimental foundation to Eschenmoser's scenario[94]. Sagi, et al. followed this initial work with a demonstration of reactions of dihydroxyfumarate with glyoxylate, glycolaldehyde, and dihydroxyacetone, to yield triose, tetruloses, and pentuloses, with remarkable efficiency (Figure 5) [95].

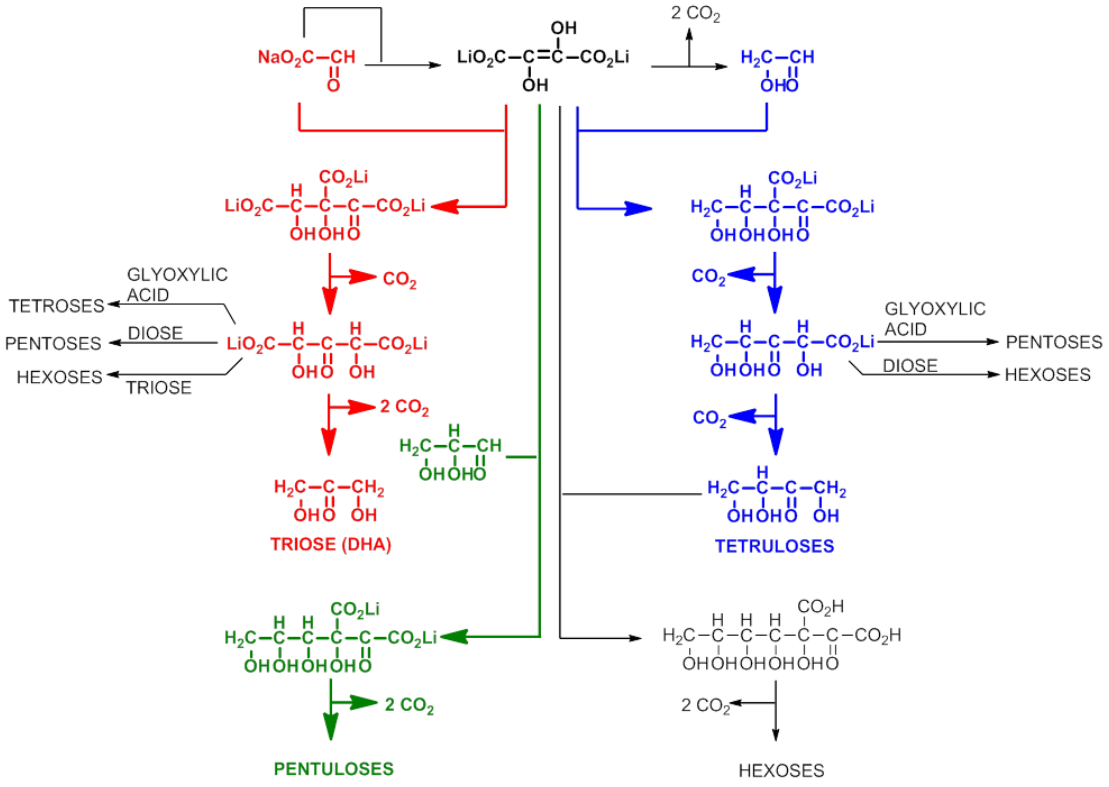


Figure 5 – Schematic of the "Glyoxylate Scenario" Mechanism for Sugar Formation as Proposed by Eschenmoser. Colored Reactions Demonstrated by Sagi et al.

These recent papers have established a foundation for prebiotic glyoxylate chemistry, however major questions remained unanswered. The first among these, reflective of the outstanding work of Sagi et al. with dihydroxyfumaric acid (DHF), is by what mechanism can glyoxylate dimerize to form DHF? Second, is DHF a necessary intermediate, or can polymerize by other mechanisms? Finally, what is the prebiotic provenance of glyoxylate? This thesis describes our findings as we set out to answer these questions.

In Chapter 2, I will describe the high pH dimerization of glyoxylate to form dihydroxyfumate. This reaction is catalyzed by cyanide in a reaction analogous to the benzoin reaction. DHF produced immediately goes on to react with an additional

equivalent of glyoxylate, producing tartrate, a molecule which could allow facile entry to the citric acid cycle. A kinetic and themodynamic model of the reaction is also provided.

In Chapter 3, I cover the production of sugars by polymerization of glyoxylate at moderately basic pH, and the production of sugar acids from DHF and aldoses at high pH. These reactions demonstrate two processes by which glyoxylate can serve as an alternative to the formose reaction for the synthesis of carbohydrates, both of which favor the production of linear sugars.

In Chapter 4, I will discuss the formation of glyoxylate from formaldehyde and cyanide (glycolonitrile) and carbon dioxide, the first prebiotically plausible synthesis of glyoxylate.

Chapter 5 contains my conclusions, and recommendations for further investigation to continue to develop the prebiotic chemistry of the glyoxylate scenario.

1.4 References

[1] DARÒS JA, et al. Viroids: an Ariadnes thread into the RNA labyrinth. EMBO reports 2006;7:593-8.

[2] COLLINS RF, et al. Self-Cleaving Circular RNA Associated with Rice Yellow Mottle Virus Is the Smallest Viroid-like RNA. Virology 1998;241:269-75.

[3] HUBER H, et al. A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. Nature 2002;417:63-7.

[4] WALKER SI, et al. Universal Sequence Replication, Reversible Polymerization and Early Functional Biopolymers: A Model for the Initiation of Prebiotic Sequence Evolution. PLoS ONE 2012;7:e34166.

[5] BUDIN I, SZOSTAK JW. Expanding roles for diverse physical phenomena during the origin of life. Annual review of biophysics 2010;39:245-63.

[6] JOYCE GF. Evolution in an RNA World. Cold Spring Harbor Symposia on Quantitative Biology 2009;74:17-23.

[7] WOESE CR. Translation: in retrospect and prospect. RNA 2001;7:1055-67.

[8] RAUCHFUSS H, MITCHELL T. The “RNA World”. Chemical Evolution and the Origin of Life: Springer Berlin Heidelberg; 2008. p. 145-80.

[9] PIZZARELLO S. Chemical Evolution and Meteorites: An Update. Orig Life Evol Biosph 2004;34:25-34.

[10] CLEAVES HJ. The prebiotic geochemistry of formaldehyde. Precambrian Research 2008;164:111-8.

[11] LOB W. Uber das Verhaltes des formamibs unter ber wirkung derstillen ent-adung: ein beitrag zue frage der flictoff-assimilton'. Chem Ber 1913;46:690.

[12] FERRIS JP, CHEN CT. Chemical evolution. XXVI. Photochemistry of methane, nitrogen, and water mixtures as a model for the atmosphere of the primitive earth. J Am Chem Soc 1975;97:2962-7.

[13] PINTO JP, et al. Photochemical production of formaldehyde in Earth's primitive atmosphere. Science 1980;210:183-4.

[14] LINDQVIST M, et al. Carbon-bearing molecules in the envelopes around oxygen-rich stars-First detection of formaldehyde in an oxygen-rich circumstellar envelope. Astronomy and Astrophysics 1992;263:183-9.

[15] AIKAWA Y, et al. Interferometric observations of formaldehyde in the protoplanetary disk around LkCa 15. Publications of the Astronomical Society of Japan 2003;55:11-5.

[16] SNYDER LE, et al. Microwave detection of interstellar formaldehyde. Physical Review Letters 1969;22:679.

[17] WICKRAMASINGHE N. Formaldehyde polymers in interstellar space. Astronomical Origins of Life: Springer; 2000. p. 111-4.

[18] SCHUTTE W, et al. Discovery of solid formaldehyde toward the protostar GL 2136: observations and laboratory simulation. Astronomy and Astrophysics 1996;309:633-47.

[19] PIRRONELLO V, et al. Formaldehyde formation in a H2O/CO2 ice mixture under irradiation by fast ions. The Astrophysical Journal 1982;262:636-40.

[20] MOORE CB, WEISSHAAR JC. Formaldehyde photochemistry. Annual Review of Physical Chemistry 1983;34:525-55.

[21] CHYBA CF, et al. Cometary Delivery of Organic Molecules to the Early Earth. Science 1990;249:366-73.

[22] CHYBA C, SAGAN C. Endogenous production, exogenous delivery and impact-shock synthesis of organic molecules: an inventory for the origins of life. 1992.

[23] ANDERS E. Pre-biotic organic matter from comets and asteroids. Nature 1989;342:255-7.

[24] AKERMARK B, et al. Photochemical, metal-promoted reduction of carbon dioxide and formaldehyde in aqueous solution. Acta Chem Scand B 1980;34:27-30.

[25] FERRIS J. Chemical Markers of Prebiotic Chemistry in Hydrothermal Systems. In: Holm NG, editor. Marine Hydrothermal Systems and the Origin of Life: Springer Netherlands; 1992. p. 109-34.

[26] INGMANSON D. Formaldehyde in hot springs. Orig Life Evol Biosph 1997;27:313-7.

[27] BUTLEROW A. Formation of monosaccharides from formaldehyde. Compt Rend 1861;53:145-67.

[28] BRESLOW R. On the mechanism of the formose reaction. Tetrahedron Letters 1959;1:22-6.

[29] SOCHA RF, et al. Homogeneously catalyzed condensation of formaldehyde to carbohydrates: VII. An overall formose reaction model. Journal of Catalysis 1981;67:207-17.

[30] KHOMENKO TI, et al. Homogeneously catalyzed formaldehyde condensation to carbohydrates: IV. Alkaline earth hydroxide catalysts used with glycolaldehyde co-catalyst. Journal of Catalysis 1976;45:356-66.

[31] PESTUNOVA O, et al. Putative mechanism of the sugar formation on prebiotic Earth initiated by UV-radiation. Advances in Space Research 2005;36:214-9.

[32] SCHWARTZ AW, DE GRAAF R. The prebiotic synthesis of carbohydrates: a reassessment. Journal of molecular evolution 1993;36:101-6.

[33] SOCHA R, et al. Autocatalysis in the formose reaction. Reaction Kinetics and Catalysis Letters 1980;14:119-28.

[34] HUSKEY WP, EPSTEIN IR. Autocatalysis and apparent bistability in the formose reaction. J Am Chem Soc 1989;111:3157-63.

[35] MIZUNO T, WEISS AH. Synthesis and utilization of formose sugars. Advances in carbohydrate chemistry and biochemistry 1974;29:173-227.

[36] LANGRIDGE R, GOMATOS PJ. The Structure of RNA. Science 1963;141:1024.

[37] KIM H-J, et al. Synthesis of Carbohydrates in Mineral-Guided Prebiotic Cycles. J Am Chem Soc 2011;133:9457-68.

[38] STOOP M, et al. Chemical Etiology of Nucleic Acid Structure: The Pentulofuranosyl Oligonucleotide Systems: The (1′→3′)-β-L-Ribulo, (4′→3′)-α-L-Xylulo, and (1′→3′)-α-L-Xylulo Nucleic Acids. Chemistry – A European Journal 2013;19:15336-45.

[39] SHIGEMASA Y, et al. Formose reactions. IV. The formose reaction in homogeneous systems and the catalytic functions of calcium ion species. Bulletin of the Chemical Society of Japan 1977;50:2138-42.

[40] ZUBAY G. Studies on the Lead-Catalyzed Synthesis of Aldopentoses. Orig Life Evol Biosph 1998;28:13-26.

[41] KHOMENKO T, et al. Homogeneously catalyzed formaldehyde condensation to carbohydrates: IV. Alkaline earth hydroxide catalysts used with glycolaldehyde co-catalyst. Journal of Catalysis 1976;45:356-66.

[42] RICARDO A, et al. Borate Minerals Stabilize Ribose. Science 2004;303:196.

[43] PETRUŠ L, et al. The Bílik Reaction. In: Stütz A, editor. Glycoscience: Springer Berlin Heidelberg; 2001. p. 15-41.

[44] STRECKER A. Ueber die künstliche Bildung der Milchsäure und einen neuen, dem Glycocoll homologen Körper. Justus Liebigs Annalen der Chemie 1850;75:27-45.

[45] BADA JL, MILLER SL. Ammonium ion concentration in the primitive ocean. Science 1968;159:423-5.

[46] SUMMERS DP. Sources and sinks for ammonia and nitrite on the early Earth and the reaction of nitrite with ammonia. Orig Life Evol Biosph 1999;29:33-46.

[47] STRIBLING R, MILLER SL. Energy yields for hydrogen cyanide and formaldehyde syntheses: the HCN and amino acid concentrations in the primitive ocean. Orig Life Evol Biosph 1987;17:261-73.

[48] RAULIN F. Prebiotic chemistry in planetary environments. In: Heidmann J, Klein M, editors. Bioastronomy The Search for Extraterrestial Life — The Exploration Broadens: Springer Berlin Heidelberg; 1991. p. 141-8.

[49] MILLER SL, VAN TRUMP JE. The Strecker synthesis in the primitive ocean. Origin of life: Springer; 1981. p. 135-41.

[50] WOESE CR. On the evolution of the genetic code. Proceedings of the National Academy of Sciences of the United States of America 1965;54:1546.

[51] WEBER AL. The sugar model: catalysis by amines and amino acid products. Orig Life Evol Biosph 2001;31:71-86.

[52] PLANKENSTEINER K, et al. Glycine and diglycine as possible catalytic factors in the prebiotic evolution of peptides. Orig Life Evol Biosph 2002;32:225-36.

[53] HUDSON JS, et al. A unified mechanism for abiotic adenine and purine synthesis in formamide. Angewandte Chemie 2012;124:5224-7.

[54] MURATA A, WADA T. Synthesis of a novel ester analog of nucleic acids bearing a serine backbone. Bioorganic & medicinal chemistry letters 2006;16:2933-6.

[55] SAWA N, et al. Synthesis and DNA-recognition behavior of a novel peptide ribonucleic acid with a serine backbone (oxa-PRNA). Tetrahedron 2010;66:344-9.

[56] NIELSEN PE, et al. Peptide nucleic acid (PNA). A DNA mimic with a peptide backbone. Bioconjugate chemistry 1994;5:3-7.

[57] NIELSEN PE. Peptide nucleic acid (PNA): a model structure for the primordial genetic material? Orig Life Evol Biosph 1993;23:323-7.

[58] AUTHOR. OBSERVATIONS OF RADIO EMISSION FROM INTERSTELLAR HYDROGEN CYANIDE. Univ. of Virginia, Charlottesville; 1971.

[59] IRVINE WM, et al. Spectroscopic evidence for interstellar ices in Comet Hyakutake. 1996.

[60] ALTMAN RS, et al. Observation of the Infrared⌫ 2 Band (CH Stretch) of Protonated Hydrogen Cyanide, HCNH+. J Chem Phys 1984;80:3911-2.

[61] TANGUY L, et al. Stratospheric profile of HCN on Titan from millimeter observations. Icarus 1990;85:43-57.

[62] SHEMANSKY DE, et al. The Cassini UVIS stellar probe of the Titan atmosphere. Science 2005;308:978-82.

[63] GEBALLE T, et al. High-resolution 3 micron spectroscopy of molecules in the mesosphere and troposphere of Titan. The Astrophysical Journal Letters 2003;583:L39.

[64] CLARKE DW, FERRIS JP. Chemical evolution on Titan: comparisons to the prebiotic Earth. Planetary and Interstellar Processes Relevant to the Origins of Life: Springer; 1997. p. 225-48.

[65] TRAINER MG, et al. Organic haze on Titan and the early Earth. Proceedings of the National Academy of Sciences 2006;103:18035-42.

[66] RAULIN F, et al. Exobiology on Titan. Journal of biological physics 1995;20:39-53.

[67] CHAMEIDES W, WALKER JC. Rates of fixation by lightning of carbon and nitrogen in possible primitive atmospheres. Origins Life Evol Biosphere 1981;11:291-302.

[68] ZAHNLE KJ. Photochemistry of methane and the formation of hydrocyanic acid (HCN) in the Earth's early atmosphere. Journal of Geophysical Research: Atmospheres (1984–2012) 1986;91:2819-34.

[69] PELTZER ET, et al. The chemical conditions on the parent body of the murchison meteorite: Some conclusions based on amino, hydroxy and dicarboxylic acids. Advances in Space Research 1984;4:69-74.

[70] FLYNN GJ, et al. The origin of organic matter in the solar system: evidence from the interplanetary dust particles. Geochimica et Cosmochimica Acta 2003;67:4791-806.

[71] LORENCAK P, et al. Iminoacetonitrile, an HCN dimer; I.R. identification in an argon matrix. Journal of the Chemical Society, Chemical Communications 1986:916-8.

[72] LOEW G, CHANG S. Energy-conformation studies of hydrogen cyanide tetramer: A prebiotic precursor. Theoret Chim Acta 1972;27:273-80.

[73] FERRIS JP, ORGEL LE. Aminomalononitrile and 4-Amino-5-cyanoimidazole in Hydrogen Cyanide Polymerization and Adenine Synthesis1. J Am Chem Soc 1965;87:4976-7.

[74] MATTHEWS C. Simultaneous synthesis of polypeptides and polynucleotides? Hydrogen cyanide polymers as prebiotic condensing agents. Orig Life Evol Biosph 1986;16:500-1.

[75] CHANG S, et al. PEPTIDE FORMATION MEDIATED BY HYDROGEN CYANIDE TETRAMER: A POSSIBLE PREBIOTIC PROCESS. Proceedings of the National Academy of Sciences 1969;64:1011-5.

[76] ORÓ J, KIMBALL A. Synthesis of purines under possible primitive earth conditions. I. Adenine from hydrogen cyanide. Archives of biochemistry and biophysics 1961;94:217-27.

[77] LOWE C, et al. SYNTHESIS OF COMPLEX ORGANIC COMPOUNDS FROM SIMPLE PRECURSORS: FORMATION OF AMINO-ACIDS, AMINO-ACID POLYMERS, FATTY ACIDS AND PURINES FROM AMMONIUM CYANIDE. Nature 1963;199:219.

[78] VOET AB, SCHWARTZ AW. Uracil synthesisvia HCN oligomerization. Origins Life Evol Biosphere 1982;12:45-9.

[79] LEVY M, et al. Production of guanine from NH4CN polymerizations. Journal of molecular evolution 1999;49:165-8.

[80] MIYAKAWA S, et al. The Cold Origin of Life: A. Implications Based On The Hydrolytic Stabilities Of Hydrogen Cyanide And Formamide. Orig Life Evol Biosph 2002;32:195-208.

[81] SALADINO R, et al. Synthesis and degradation of nucleobases and nucleic acids by formamide in the presence of montmorillonites. ChemBioChem 2004;5:1558-66.

[82] MOUTOU G, et al. Equilibrium of α-aminoacetonitrile formation from formaldehyde, hydrogen cyanide and ammonia in aqueous solution: Industrial and prebiotic significance. Journal of Physical Organic Chemistry 1995;8:721-30.

[83] RITSON D, SUTHERLAND JD. Prebiotic synthesis of simple sugars by photoredox systems chemistry. Nature chemistry 2012;4:895-9.

[84] ARRHENIUS G, et al. Glycolonitrile Oligomerization: Structure of Isolated Oxazolines, Potential Heterocycles on the Early Earth. The Journal of Organic Chemistry 1997;62:5522-5.

[85] SCHWARTZ AW, BAKKER C. Was adenine the first purine? Science 1989;245:1102-4.

[86] BRAUNSHTEIN. International Symposium on the Origin of Life on the Earth Proceedings. London: Pergamon Press; 1959.

[87] METZLER DE, et al. A general mechanism for vitamin B6-catalyzed reactions1. J Am Chem Soc 1954;76:648-52.

[88] MELÉNDEZ-HEVIA E, et al. The puzzle of the Krebs citric acid cycle: Assembling the pieces of chemically feasible reactions, and opportunism in the design of metabolic pathways during evolution. Journal of Molecular Evolution 1996;43:293-303.

[89] FOX SW, et al. Photochemical synthesis of ATP: protomembranes and protometabolism. Light Transducing Membranes: Structure, Function, Evolution (ed) Deamer, DW 1978:61.

[90] MOROWITZ H, et al. The synthesis of glutamic acid in the absence of enzymes: Implications for biogenesis. Orig Life Evol Biosph 1995;25:395-9.

[91] YANAGAWA H, et al. Novel formation of α-amino acid from α-oxo acids and ammonia in an aqueous medium. Origins Life Evol Biosphere 1984;14:163-9.

[92] BEAN H, et al. Glyoxylate as a Backbone Linkage for a Prebiotic Ancestor of RNA. Origins of Life and Evolution of Biospheres 2006;36:39-63.

[93] ESCHENMOSER A. The search for the chemistry of life's origin. Tetrahedron 2007;63:12821-43.

[94] NAIDU SAGI V, et al. Diastereoselective Self-Condensation of Dihydroxyfumaric Acid in Water: Potential Route to Sugars. Angewandte Chemie International Edition 2011;50:8127-30.

[95] SAGI VN, et al. Exploratory Experiments on the Chemistry of the “Glyoxylate Scenario”: Formation of Ketosugars from Dihydroxyfumarate. J Am Chem Soc 2012;134:3577-89.

CHAPTER 2:
DIMERIZATION OF GLYOXYLATE AND PRODUCTION OF CITRIC ACID CYCLE INTERMEDIATES

Despite the success enjoyed by Sagi and coworkers, the glyoxylate scenario still faced a critical challenge in that dimerization of glyoxylate to form dihydroxyfumarate had not been experimentally demonstrated. The nearest thing to a prebiotic synthesis of dihydroxyfumarate was demonstrated by Saladino et al. wherein they demonstrated photocatalytic hydrolysis of the cyanide tetramer over titanium dioxide[1], a mineral not commonly thought to have been prebiotically available. Consequently, we set out to demonstrate the synthesis of DHF, believing that accounting for its origin would cement the glyoxylate scenario as a plausible prebiotic scenario. To this end we chose to investigate reaction of glyoxylate and cyanide as an umpolung catalyst (a possibility suggested by Eschenmoser[2]), which we have named the glyoxoin reaction, depicted in Figure 6.

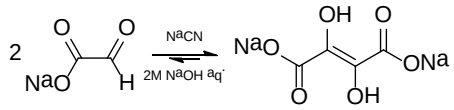


Figure 6 – Simplified depiction of the proposed glyoxoin reaction.

We have named the glyoxoin reaction thusly to draw analogy to the benzoin reaction [3] wherein the aldehyde functionality of benzaldehyde is activated by cyanide, in the presence of base, allowing dimer formation. We proposed that the glyoxoin reaction may function analogously and provide a catalytic, prebiotically plausible route to dihydroxyfumarate formation.

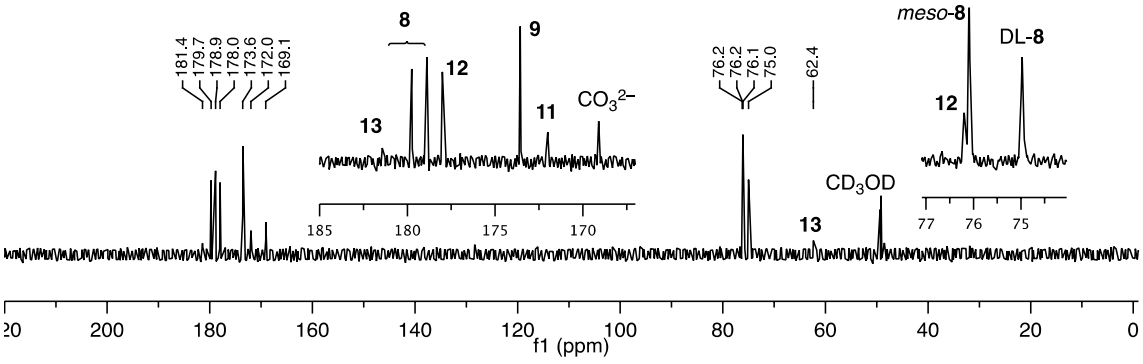


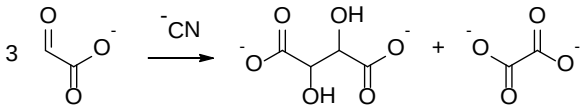
Figure 7 – Typical ^{13}C NMR of the glyoxoin reaction. Reaction of 1 M sodium glyoxylate (182/94 ppm – completely consumed) with 0.1 M cyanide (167ppm) in aqueous 2 M NaOH (room temperature, 1 hour) produces *meso*-tartrate 8 (178.9/76.1ppm) and D,L-tartrates 8 (179.8/75.0 ppm). Signals for carbonate (169.1 ppm), oxalate 9 (173.6 ppm), formate 11 (172ppm), tartronate 12 (178.0/76.2ppm) and glycolate 13 (trace, 181.4/62.4) were observed. CD_3OD was used as external standard (49.15 ppm).

The (homogeneous) glyoxoin reaction of glyoxylate (1.0 M) with a catalytic amount of NaCN (0.1 M) in aqueous medium at room temperature in 2.0 M NaOH (pH ≈ 14) showed complete consumption of glyoxylate (by 13C NMR after 28 h); no signals corresponding to dihydroxyfumaric acid were exhibited, contrary to our expectations for this reaction. Instead, as seen in Figure 7, meso, and D/L tartrates **8**, the reduced forms of dihydroxyfumarate **5,** were formed in high yield, exceeding 80% in certain cases, see Table 1. Additional products included oxalate **9**, formate **10,** tartronate **12**, glycolate **13,** and carbonate. The collected NMR is remarkably clean after only one hour of reaction time with no trace of starting material remaining (expected peaks of 94 and 182 ppm). All of the identified products were confirmed by spiking of the NMR with authentic materials (Figure 8). Spectral shifts identified from this spiking experiment are cataloged in Table 2. To positively identify tartrate **8** in the reaction mixture, the methods of Hearon, et al. were used to successfully separate and crystallize the dicalcium tartrate salt[4]. Both NMR and LCMS were used to confirm the identity of this recovered

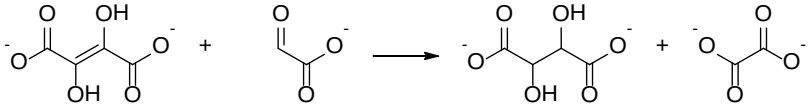
dicalcium tartrate salt (Figure 9). In all experiments meso-tartrate **8** was observed to form in equal or greater concentration, as compared to the DL-tartrates **8**. This excess appears to be kinetically governed, as no interconversion of D-tartrate to meso tartrate (nor formation of any other product) was observed when authentic D-tartrate was subjected to reaction conditions for 8 days (Figure 10).

Table 1 – Yields of Glyoxoin Reaction and Aldol Reaction-Fragmentation Reaction. Yields were calculated by comparing quantitative carbon integrals of each species versus the total carbon integral and the known concentration of carbon in the sample.

Glyoxoin Reaction Yield Based on:



Aldol Reaction-Fragmentation Yield Based on:



	Glyoxoin Reaction: 0.75 M Glyoxylate 0.075 M NaCN 2 M NaOH 1 Day		Aldol Reaction-Fragmentation: 0.25 M Glyoxylate 0.25 M Dihydroxyfumarate 1 M NaOH 1 M LiOH 1 Day	
	Yield (% of expected)	% of total Carbon	Yield (% of expected)	% of total Carbon
Oxalate	73.0%	23.2%	82.7%	27.6%
Total Tartrates	80.8%	51.3%	69.0%	46.0%
(D/L-Tartrates)	(39.6%)	(25.1%)	(31.0%)	(20.6%)
(Meso-Tartrate)	(41.3%)	(26.2%)	(38.0%)	(25.4%)
Tartronate	-	12.6%	-	6.7%
Carbonate	-	6.4%	-	2.8%
Formate	-	2.2%	-	1.0%
Glycolate	-	1.1%	-	4.7%
Cyanide	-	3.2%	-	-
	Conversion	% of total Carbon	Conversion	% of total Carbon
Glyoxylate	100%	0%	66.47%	11.2%
DHF	-	-	100%	0%

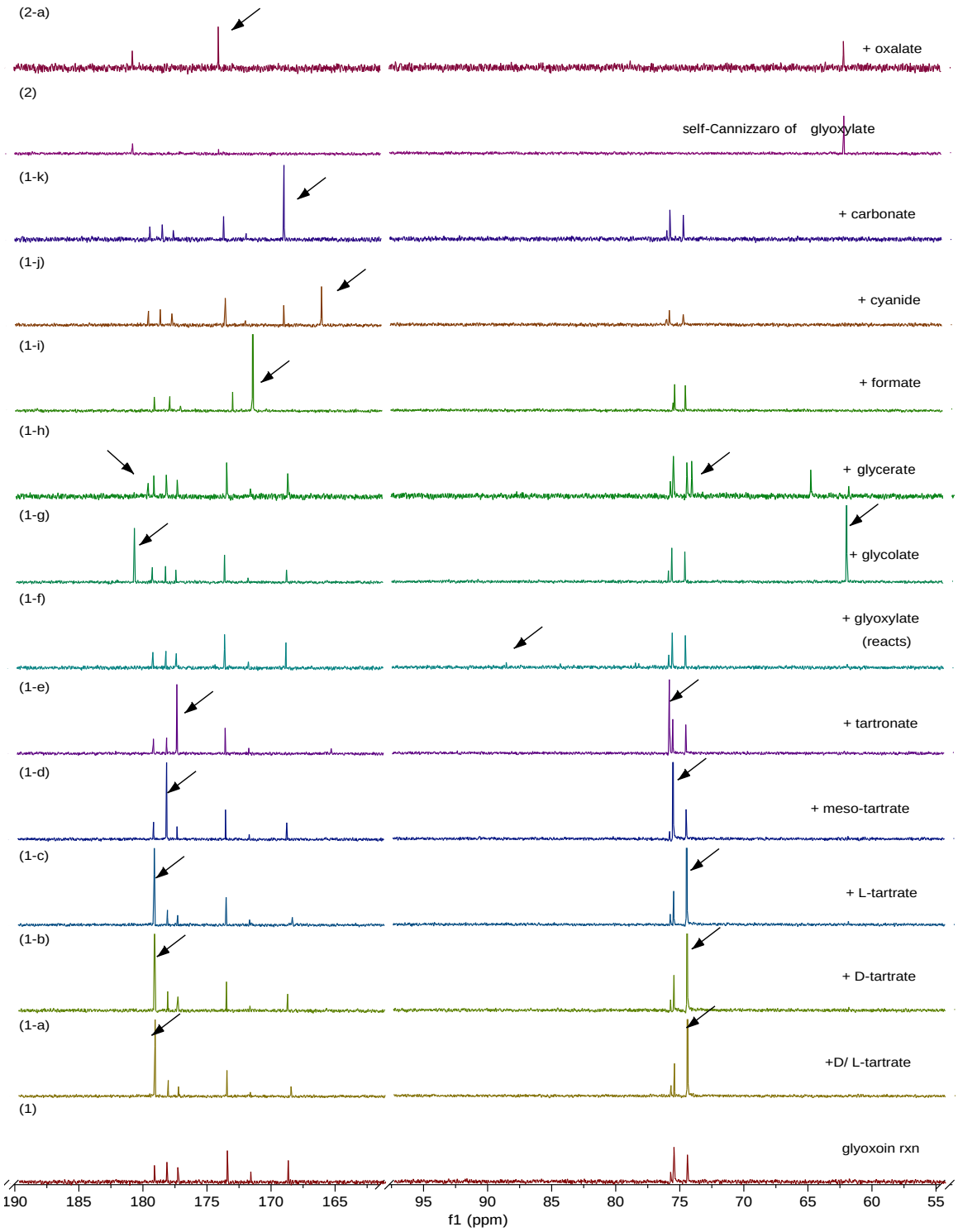


Figure 8 – Glyoxoin and Cannizarro Reactions Spiked with Authentic Standards to Confirm Peak Identities. (1) Reaction of 1M glyoxylate and 0.1M NaCN in 2 M NaOH after 1 day spiked with (a) D/L-tartrate (b) D-Tartrate (c) L-Tartrate (d) *meso*-Tartrate (e) Tartronate (f) Glyoxylate (reacts immediately) (g) Glycolate (h) Glycerate (i) Formate (j) Cyanide (k) Carbonate. (2) Self-Cannizzaro products of 1M glyoxylate in 2M NaOH after 1 day, spiked with (2a) Oxalate.

Table 2 – ^{13}C NMR chemical shifts (in H_2O with CD_3OD as external standard) of all the species observed in the glyoxoin and the DHF/glyoxylate reaction. Peaks may exhibit some shift with changes in concentration and prevailing counter ion.

Species	Carbon	PPM
Oxalate	COO-	173.7
D/L-Tartrates	COO-	179.7
D/L-Tartrates	HCOH	74.9
meso-Tartrates	COO-	178.8
meso-Tartrates	HCOH	76.0
Formate	HCOO-	172.0
Glycolate	COO-	181.5
Glycolate	H_2COH	62.1
Glyoxylate	COO-	182.8
Glyoxylate	$HC(OH)_2$	93.7
Carbonate	-OCOO-	169.1
Tartronate	COO-	178.0
Tartronate	HCOH	76.2
Glyoxylate Cyanohydrin	COO-	178.0
Glyoxylate Cyanohydrin	CN	124.5
Glyoxylate Cyanohydrin	HCOH	97.1
Dihydroxyfumarate	COO-	176.6
Dihydroxyfumarate	COH	139.6

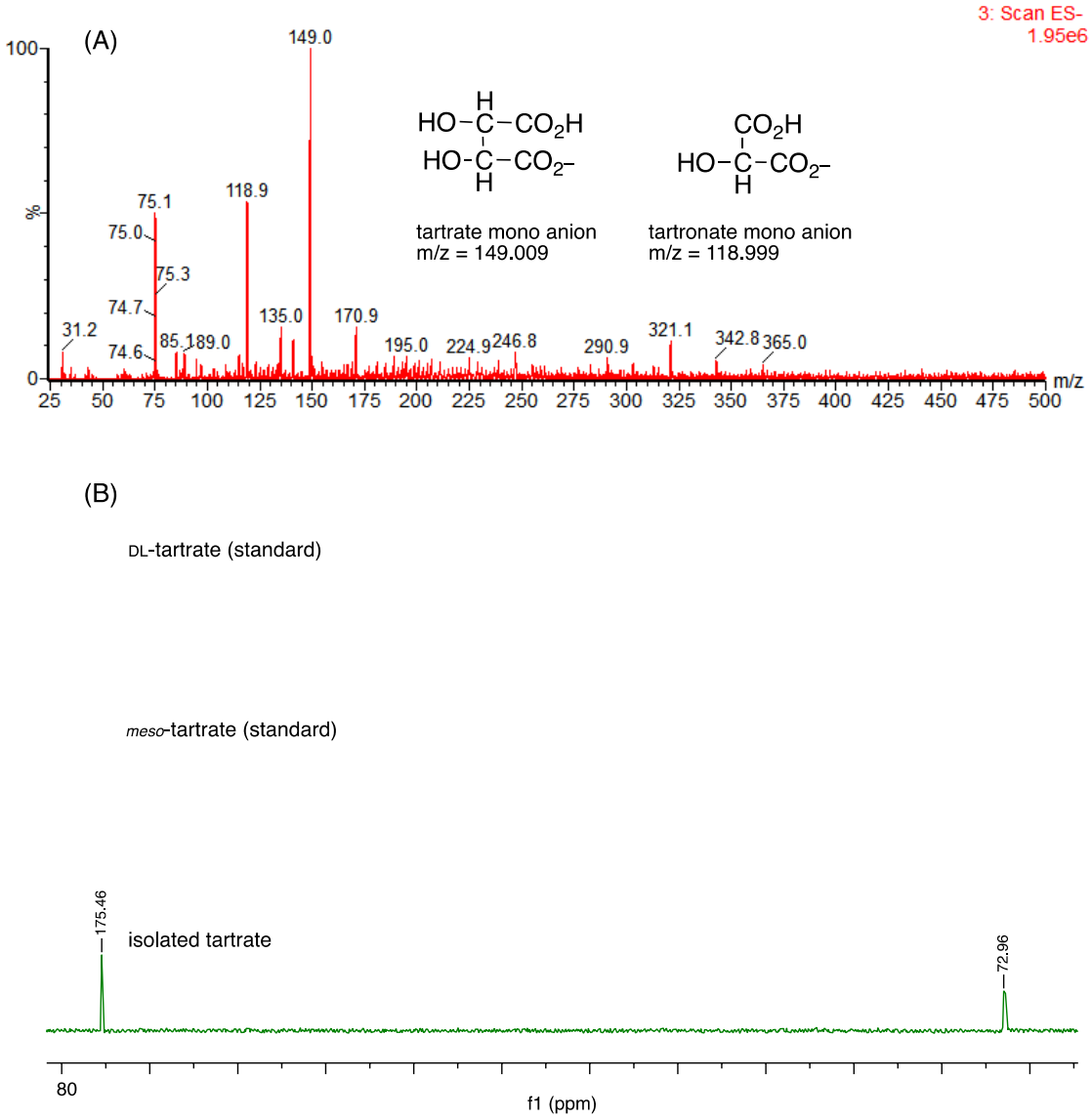


Figure 9 – Spectral data of isolated tartrate from the glyoxoin reaction. (A) "Direct inject" LCMS spectrum of isolated tartrate from the glyoxoin reaction. Data was obtained from negative mode electrospray ionization. The peak at 75 is from the mobile phase. (B) ^{13}C NMR comparing authentic D/L- and *meso*-tartrates with (re-acidified) material isolated from glyoxoin reaction. Presence of largely D/L-tartrate in the isolated material suggests that *meso*-tartrate was lost during washing or precipitated during removal of oxalate.

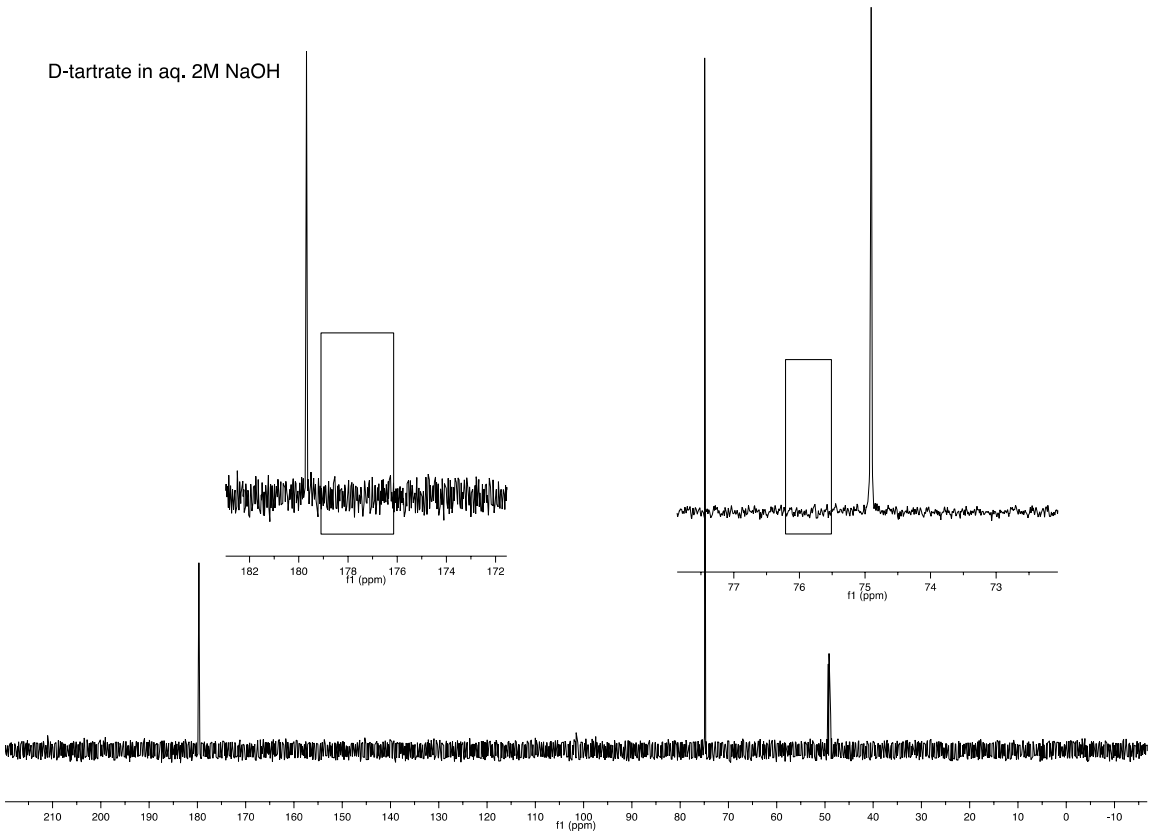


Figure 10 – ^{13}C NMR spectra of authentic D-tartrate in 2 M NaOH after 8 days at room temperature. No signals corresponding to *meso*-tartrate could be seen. This indicates that racemization of D-tartrate (to form *meso*-tartrate) does not occur under the reaction conditions.

The glyoxoin reaction also produced tartrates **8** at lower concentrations (0.01 M of **1** and 0.002 M CN^-) and at lower temperatures (4°C). The reaction is, in fact, so resilient that even under heterogeneous conditions, (when insoluble lithium glyoxylate was substituted in place of sodium glyoxylate) tartrates were still produced (Figure 11). However, when the pH was lowered (pH ≈ 9), tartrates were not observed indicating that high pH was crucial. When a control reaction (at pH ≈ 14) was conducted by omitting the cyanide ion, no tartrates were formed; only the disproportionation of glyoxylate to glycolate **13** and oxalate **9** was observed, proceeding through the well-known Cannizzaro reaction (Figure 14) [5, 6].

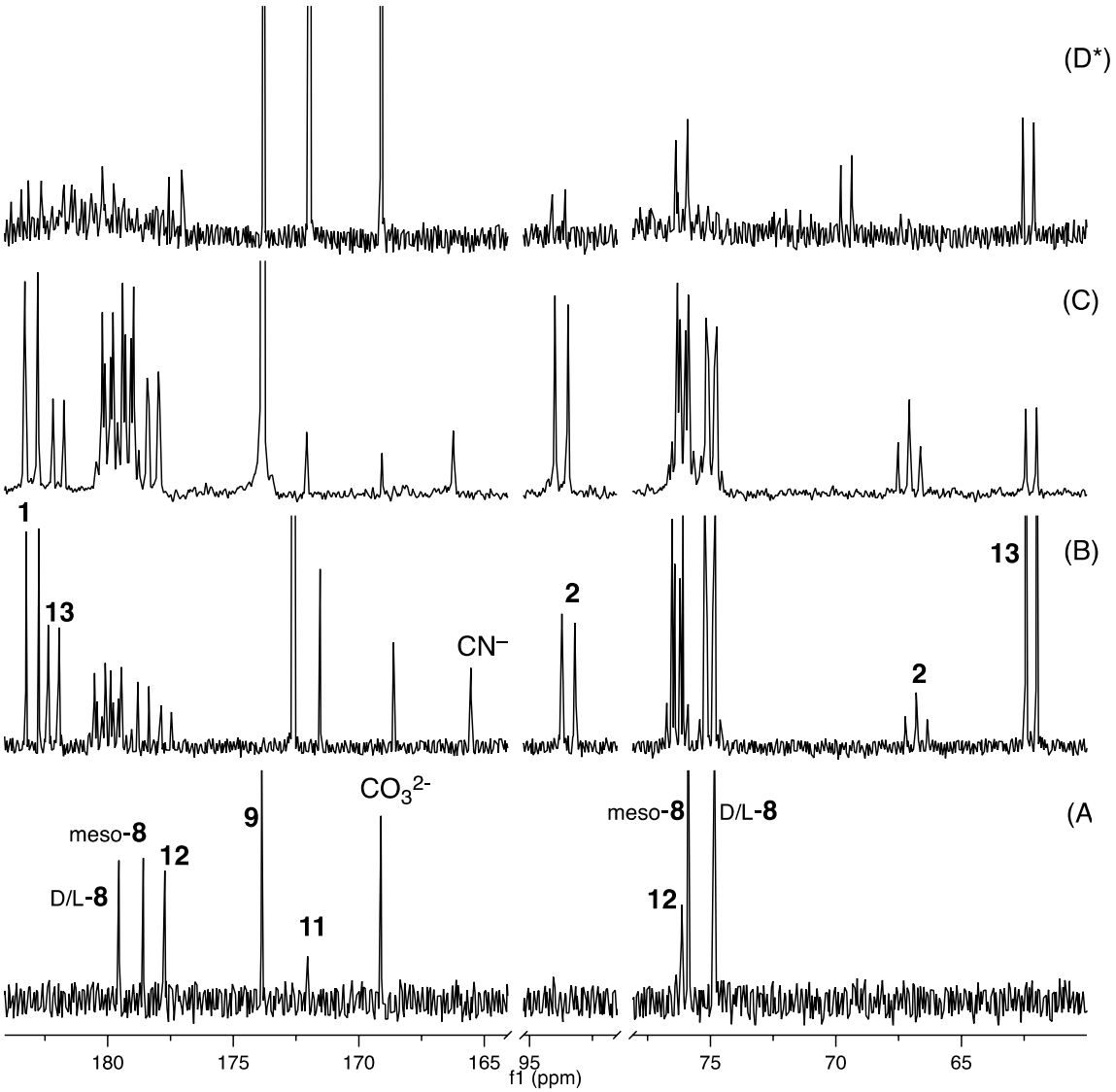


Figure 11 – ^{13}C NMR spectra of the glyoxoin reaction at varying temperatures and concentrations (Peaks are not quantitative and are scaled to oxalate, and vertically truncated to show smaller peaks): (A) 1 M glyoxylate, 0.1 M NaCN, 2 M NaOH, room temperature, 1 h. (B) Heterogeneous reaction: 1 mmol/mL ^{13}C-di-labeled glyoxylate, 0.01M ^{13}C-labeled NaCN, 2 M LiOH, room temperature, 1d. (C) 0.1 M ^{13}C-di-labeled glyoxylate, 0.01M NaCN, 2 M NaOH, 4 °C, 72 h. (D) 0.010 M ^{13}C-di-labeled glyoxylate, 0.002 M NaCN, 2 M NaOH, room temperature, 2 weeks. Glyoxylate 1; glycolate 13; DL-tartrates 8; *meso*-tartrate 8; tartronate 12; oxalate 9; formate 11; carbonate; cyanide; cyanohydrin of glyoxylate 2.

D*: Not all peaks in D have been identified; however tartrate does appear to be produced.

A comparison of Figure 14 A and C yields an interesting insight into the potential origin of tartaric acid in this system. In the Cannizzaro reaction of glyoxylate **1**, oxalate **9** and glycolate **13** are formed in equal parts, in the glyoxoin reaction however, an abundance of oxalate **9** is formed with no visible glycolate **13** formation. As oxalate **9** is the oxidation product of glyoxylate **1** , and tartrate **8** is reduced dihydroxyfumarate **5**, this gave an indication that DHF may have been produced and subsequently reduced in the glyoxoin reaction.

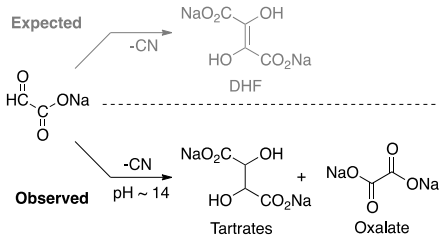


Figure 12 – Cyanide mediated dimerization of glyoxylate leads to (predominantly) tartrates and oxalate.

2.1 Reaction Mechanism

2.1.1 Was Tartrate Produced from Dihydroxyfumarate?

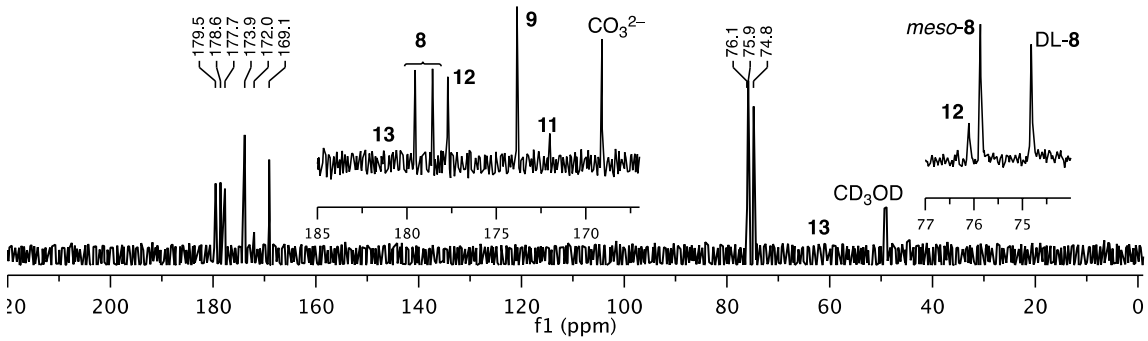


Figure 13 – Typical ^{13}C NMR of the homogeneous DHF/glyoxylate (without cyanide) reaction. Reaction of 0.25 M glyoxylic acid with 0.25 M DHF in aqueous 1 M NaOH and 1M LiOH (room temperature, 1 hour). For NMR details see caption of Figure 1

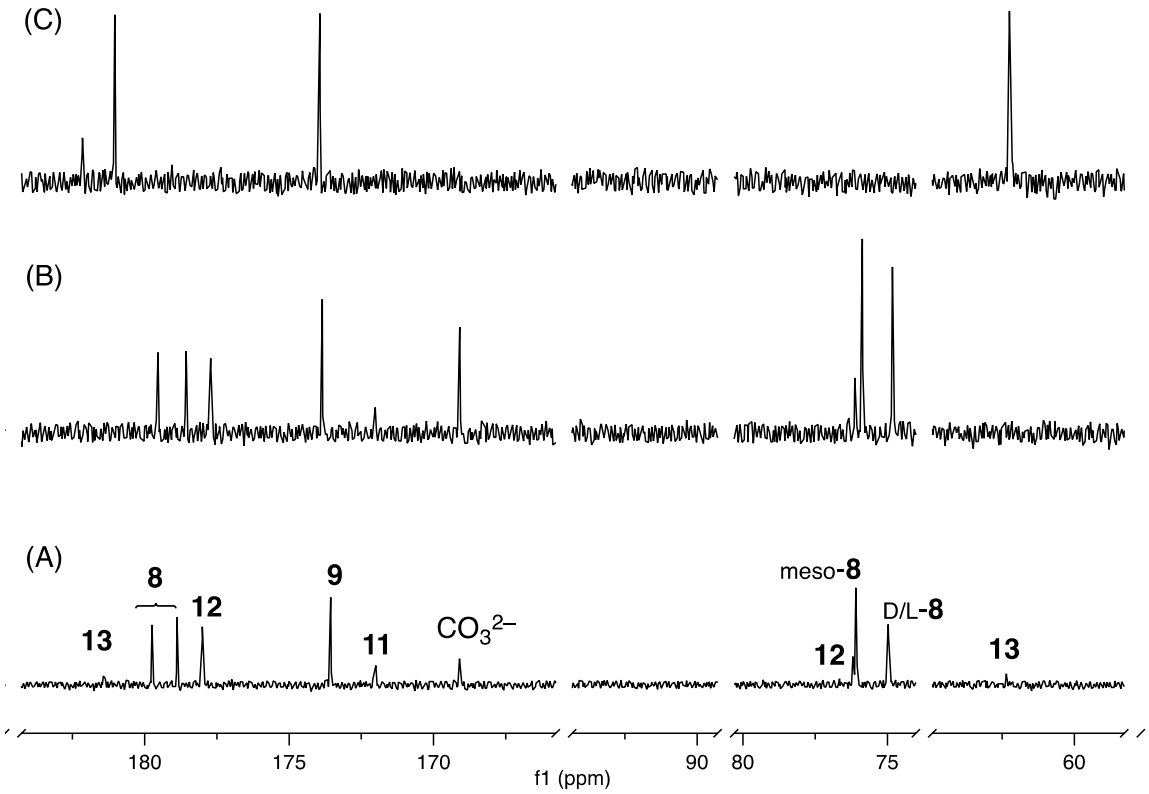


Figure 14 – Comparison of ^{13}C-NMR spectra of Glyoxoin, DHF/Glyoxylate, and Cannizzaro Reactions. (A) The homogeneous DHF/glyoxylate reaction: 0.25 M glyoxylate, 0.25 M DHF, 1 M NaOH, 1 M LiOH, room temperature, 1 h, and (B) the glyoxoin reaction: 1M glyoxylate, 0.1 M NaCN, 2 M NaOH, room temperature, 1 h. (C) Glyoxylate control: 1M glyoxylate, 2M NaOH, room temperature, 24 hr. Glyoxylate 1; glycolate 13; DL-tartrate 8; *meso*-tartrate 8; tartronate 12; oxalate 9; formate 11 and carbonate.

To test whether tartrate was produced in a secondary reaction by formal reduction of DHF (which could also lead to observed oxalate), dihydroxyfumaric acid (5mmol)was reacted with glyoxylic acid (5 mmol) and sodium cyanide (1 mmol) in 2M sodium hydroxide. This reaction was initially conducted heterogeneously, due to the sparing solubility of the disodium dihydroxyfumarate salt. A vigorous bubbling was observed in this heterogeneous reaction mixture for approximately two hours, at which point a ^{13}C NMR spectrum of the supernatant showed the production of tartrates **8** along with carbonate, oxalate **9**, formate **10**, tartronate **12**, and glycolate **13**. More revealingly, when the DHF/glyoxylate reaction was repeated, omitting the cyanide, tartrates **8** were still

formed. When the reaction of 0.25 M of DHF **5** with 0.25 M glyoxylic acid (in the absence of cyanide) was repeatedin a mixture of 1.0 M NaOH and 1.0 M LiOH, the reaction mixture became homogeneous (similar to the glyoxoin reaction). ^{13}C-NMR spectrum, after one hour, showed production of tartrates **8** ‡‡ with complete glyoxylate conversion and little carbonate formation (Figure 13); the ratios of the accompanying product peaks were similar to the glyoxoin reaction (Figure 14 B) with higher proportion of *meso*-tartrate (Table 1). Interestingly, in the heterogeneous case, the ^{13}C NMR spectrum of the reaction was also nearly identical to the glyoxoin reaction with the exception of a much more intense carbonate peak (Figure S6, B). In both the heterogeneous and homogeneous case, these results demonstrated that DHF could, by itself, mediate the transformation of glyoxylate to tartrates (Figure 15).

This production of tartrates from the reaction of DHF and glyoxylate should be compared to the work of Sagi, et al. wherein they reacted dihydroxyfumarate with glyoxylate in lithium hydroxide at pH 7-8. This reaction ultimately produced dihydroxyacetone by way of a six carbon, tricarboxylic acid intermediate, and several decarboxylations. While the high concentration of carbonate produced is indicative of decarboxylations occurring, no evidence was seen of any of the intermediates or products observed in Sagi's work. This suggests that the high concentration of sodium hydroxide is critical in determining the resultant product of this reaction.

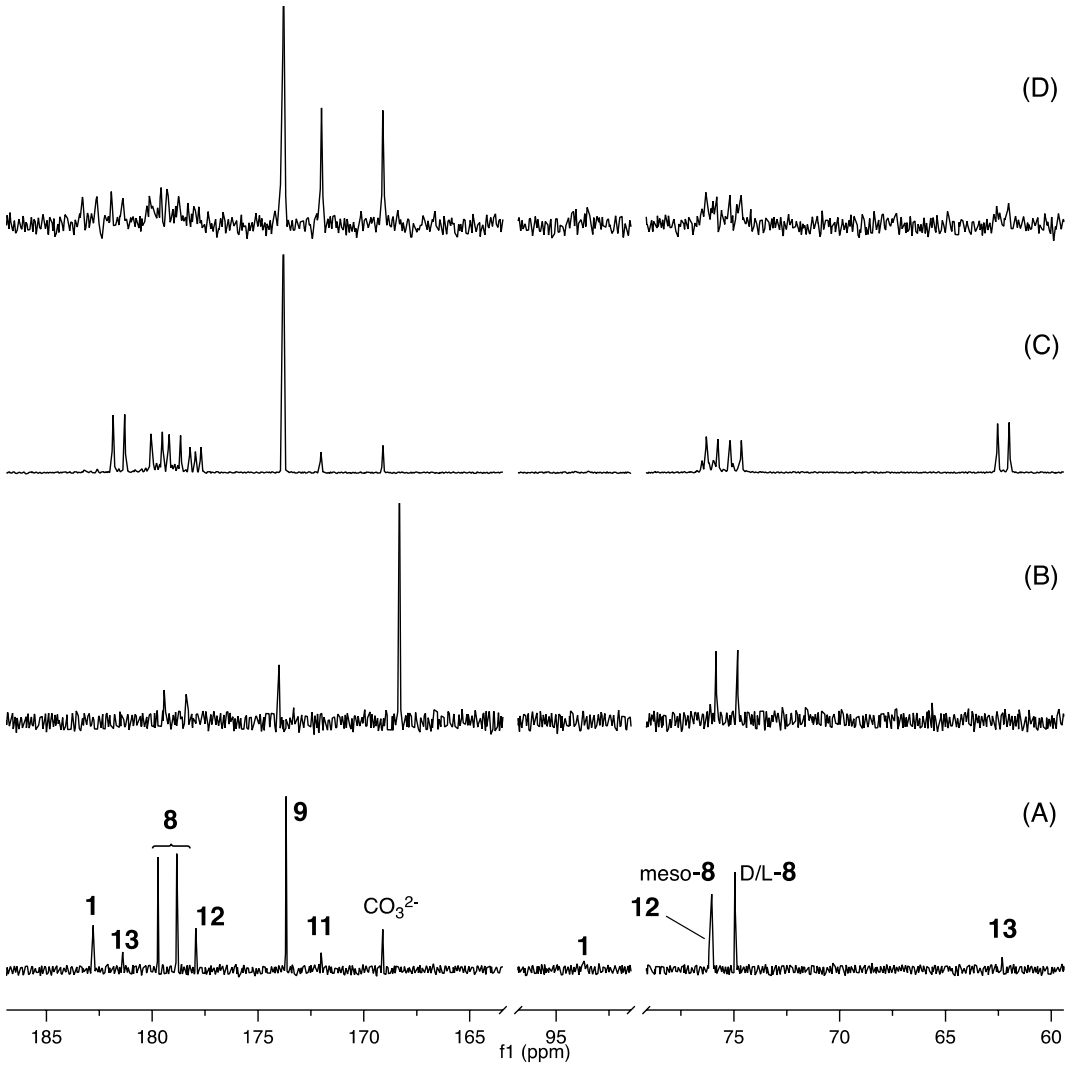


Figure 15 – ^{13}C NMR spectra documenting the DHF/glyoxylate reaction at different concentrations. (A) Homogeneous reaction: 0.25 M glyoxylate, 0.25 M DHF, 1 M NaOH, 1 M LiOH, room temperature, 2 hr. (B) Heterogeneous reaction: 0.1 M glyoxylate, 0.1mmol/mL, DHF, 2 M NaOH, room temperature, 2 weeks. (C) Heterogeneous reaction: 0.1M ^{13}C-di-labeled GA, 0.1mmol/mL, DHF, 2M NaOH, room temperature, 2 weeks. (D) Homogeneous reaction: 0.005 M ^{13}C-di-labeled glyoxylate, .005 M DHF, 2 M NaOH, room temperature, 5 weeks. Glyoxylate (**1**); glycolate (**13**); D/L-tartrates (**8**); *meso*-tartrate (**8**); tartronate (**12**); oxalate (**9**); formate (**11**).

Based upon these observations, we postulate that reduction of dihydroxyfumarate to tartrate takes place by one of two pathways(Figure 16). The first pathway is a cross-Cannizzaro pathway[7] wherein the deprotonated hydrate of glyoxylate acts as a hydride donor to the keto- form of dihydroxyfumarate. The second possible pathway is an aldol-retroaldol reaction pathway. In this pathway the DHF attacks the glyoxylate carbonyl,

resulting in the same six carbon, tri-carboxylic acid intermediate invoked by Sagi et al. Diverging from the Sagi reaction pathway, this intermediate is attacked at its carbonyl by a hydroxyl group, and subsequently fragments, yielding oxalate **9** and a tartrate **8** carbanion which will abstract a proton from water. Tartrates could have been additionally produced from the base catalyzed condensation reaction of glycolate **13** (a product of the Cannizzaro reaction) with glyoxylate **1**. This pathway was ruled out, since tartrates were observed neither in the glyoxylate control reaction discussed above nor when glyoxylate and glycolate were combined under identical reaction conditions.

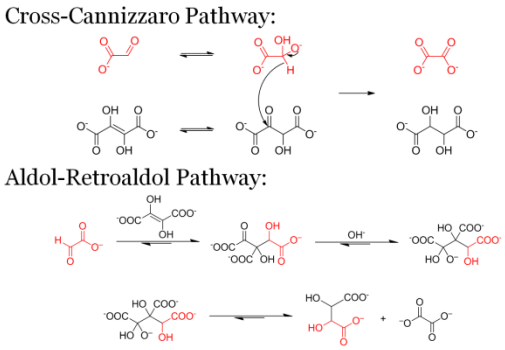


Figure 16 – Proposed Reduction Pathways, Fate of Glyoxylate is Shown in Red.

2.1.2 Isotopically Labeled Reactions

To differentiate between the two possible reactions illustrated in Figure 16 we performed a reaction of DHF and ^{13}C dilabeled glyoxylate in 2M sodium hydroxide. As before in sodium hydroxide, this reaction was initially hetero generous, but due to lower concentrations, became homogeneous as the reaction progressed. If the cross-Cannizzaro hydride transfer was the sole pathway, only the formation of ^{13}C labeled oxalate **9*** would be expected. If the fragmentation of the six-carbon tricarboxylate **6*** was the only pathway, only signals corresponding to labeled tartrates **8*** would be observed. However, when ^{13}C di-labeled glyoxylic acid (0.1 M) was reacted with DHF (0.1 M) in 1.0 M

NaOH and 1.0 M LiOH (homogeneous reaction), the ^{13}C NMR spectrum showed that both labeled tartrates **8*** and oxalate **9*** were formed.

While the presence of labeled tartrates **8*** indicates that an aldol reaction and fragmentation reaction must take place, the self-Cannizzaro reaction of glyoxylate (as opposed to cross Cannizzaro which would yield tartrate) is a second pathway to the formation of labeled oxalate **9***. Therefore in order to determine whether the cross-Cannizzaro reaction is taking place, integrals of the oxalate and glycolate peaks in the product NMR must be compared, this reveals that in fact, there is some 50% less labeled oxalate than labeled glycolate in the reaction mixture (Figure 17 C). This excess of glycolate indicates two things. First there is likely some side reaction pathway that results in the production of labeled glycolate **13***. Second, if the hydride transfer reaction is taking place, it is doing so in lower yield than this glycolate producing side reaction. Therefore, although the cross-Cannizzaro reaction cannot be ruled out in its entirety, this experiment does indicate the aldol-retroaldol pathway to be dominant.

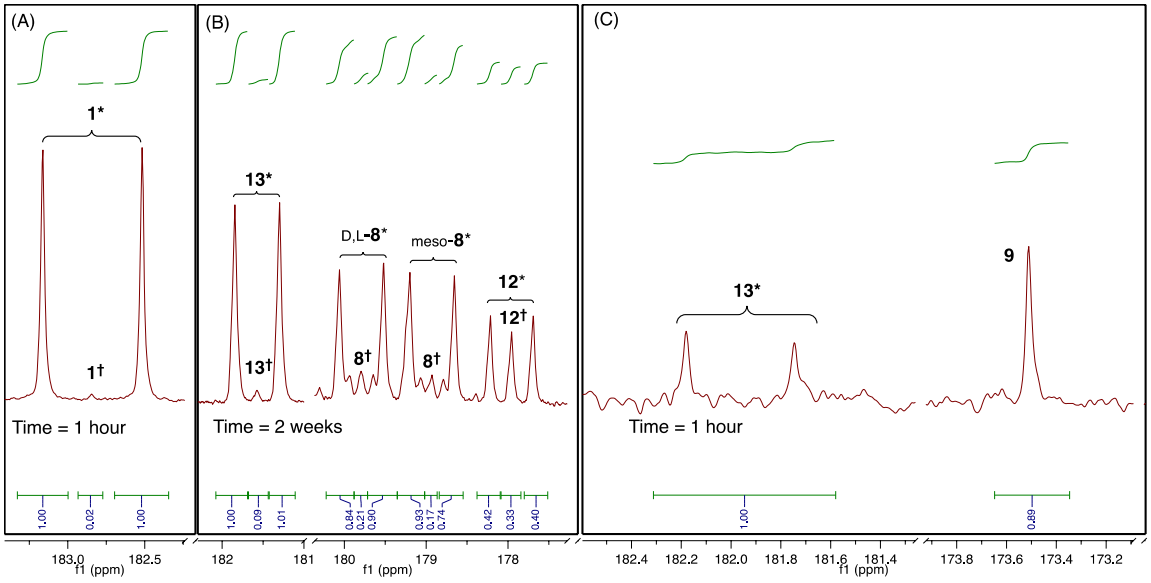


Figure 17 – Illustrative peaks from the ^{13}C NMR spectrum of reaction of the heterogeneous DHF/glyoxylate reaction. (A) Comparison of initial glyoxylate 1* integral versus singly labeled 1† integral; 0.1 M ^{13}C glyoxylate and 0.1 M DHF in 2 M NaOH at room temperature, 1 hour. (B) Comparisons of glycolate 13*/13†, tartrates 8*/8†, and tartronate 12*/12†; 0.1 M ^{13}C glyoxylate and 0.1 M DHF in 2 M NaOH at room temperature, 2 weeks. Higher ratios of singly labeled peaks in (A) versus (B) indicate breaking of the glyoxylate C-C bond, indicative of the benzoin type rearrangement. (C) Comparison of glycolate and oxalate integrals; 0.1 M ^{13}C glyoxylate and 0.1 M DHF in 1 M NaOH and 1 M LiOH at room temperature after 1 hour. Oxalate integral must be halved to account for symmetry. This results in a ratio of labeled glycolate to labeled oxalate of 2:1 indicating that labeled glycolate is produced outside of the self-Cannizzaro reaction at a faster rate than labeled oxalate. This labeling strategy results in 98.9% of the ^{13}C in the mixture originating from the glyoxylate.

A separate reaction using unlabeled glyoxylate and ^{13}C labeled NaCN (0.01 M), clearly revealed the glyoxylate cyanohydrin **2** at ≈ 125 ppm an otherwise empty spectral region; as the reaction progressed, a second peak appeared at ≈ 126 ppm (attributed to the DHF cyanohydrin **4**) persisting for approximately 4 hours before disappearing (Figure S8).

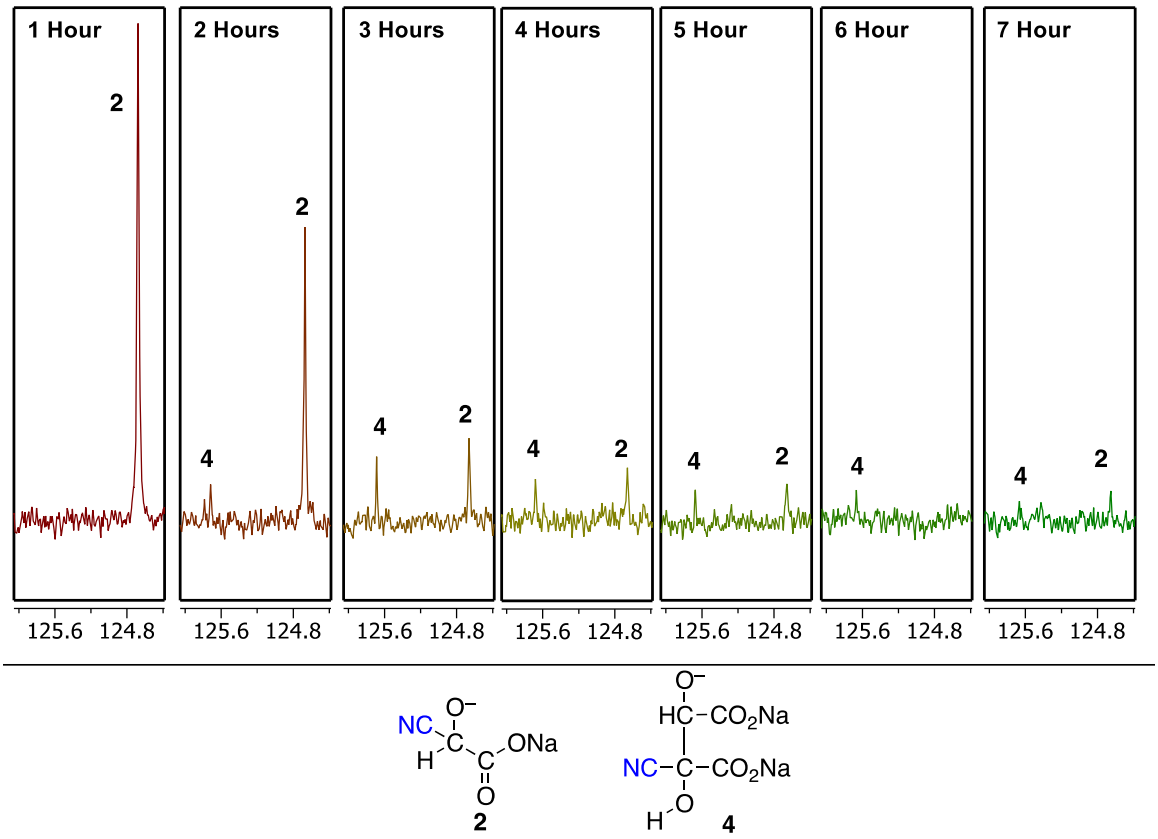


Figure 18 – ^{13}C NMR spectrum of reaction of 1 M glyoxylate and 0.01 M ^{13}C-labeled NaCN in 2 M NaOH at room temperature. Shown are expanded region for the signals of the cyano carbons (blue) of the glyoxylate cyanohydrin **2** (124.7 ppm) and the DHF cyanohydrin **4** (125.7 ppm) as the reaction progresses.

In the heterogeneous reaction (in 2M aq. NaOH) of labeled glyoxylate with DHF, a second pathway was suggested by the formation of significant amount (~10%) of singly labeled tartrates **8**† (Figure 17-B), along with singly labeled tartronate **12**†. Interestingly, the (single) labeling occurs only at the carboxylate moiety of tartrates **8**† and tartronate **12**† as evidenced by a lack of splitting of the carboxylate signal. Some of this could be explained by incomplete labeling of the starting material; however, the starting material contains less than 1% of singly labeled material (by ^{13}C NMR, Figure 17-A). This suggested that the carbon-carbon bond in glyoxylate is being broken during the course of the reaction.

2.1.3 Glyoxoin Mechanism

Having determined the aldol-retro aldol pathway to be the dominant pathway in reduction of dihydroxyfumarate to tartrate, we incorporated this mechanism into our original glyoxoin reaction to provide an overall pathway for the production of tartrate Based on the above observations, we propose an overall mechanism (Figure 19) which accounts for the bulk of the tartrates **8**, and oxalate **9**. In this mechanism, cyanide adds to glyoxylate **1** to form the glyoxylate cyanohydrin **2**. The deprotonated glyoxylate cyanohydrin **3** then reacts with an additional molecule of glyoxylate to form the cyanohydrin-adduct of DHF **4**, which is then converted to the keto form of DHF **5** (which tautomerizes to the typical enol form). DHF, thus formed, can react via an aldol reaction with an additional molecule of glyoxylate (^{13}C labeling shown in red) to yield a six-carbon tricarboxylate intermediate **6***, which can (under high pH) rearrange to **7*** via a cyanide mediated retro-aldol pathway. These intermediates **6***/**7*** can be attacked at the carbonyl by a free hydroxide anion (under high pH) and undergo fragmentation to yield tartrates **8**/**8*** and oxalate **9/9***. This mechanism accounts for the primary pathway to tartrates, in which oxalate and tartrates are formed in a 1:1 ratio.

Additional side products are formed by pathways (**Figure 20**) that include a series of benzoin-type rearrangements,[8] which are possible under these high alkaline conditions,[9, 10] and which also account for the formation of singly labeled products. In these pathways the tri-carboxylate intermediate **6**(**6***)/**7**(**7***) can undergo benzoin type

rearrangements to a six-carbon aldehyde intermediate which fragments to yield singly labeled tartrates and formate.

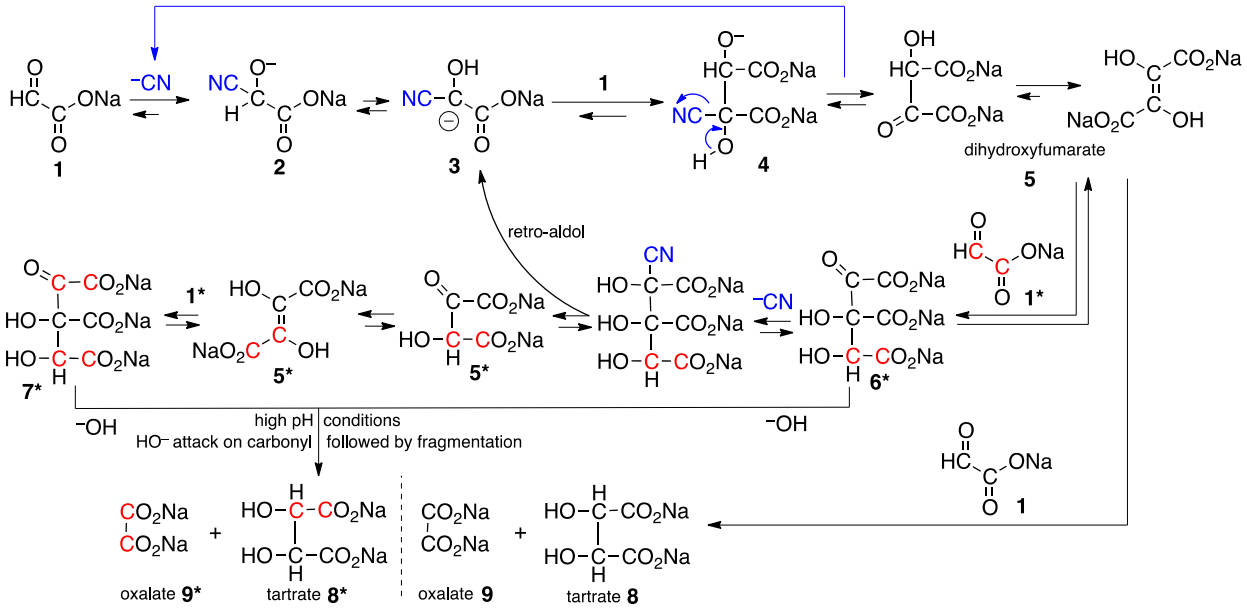


Figure 19 – Major pathway leading to tartrates and oxalate from glyoxylate: Cyanide (shown in blue) adds to glyoxylate **1** to form the glyoxylate cyanohydrin **2**. The deprotonated glyoxylate cyanohydrin **3** then reacts with an additional molecule of glyoxylate to form the cyanohydrin-adduct of DHF **4**, which is then converted to the keto form of DHF **5** (which tautomerizes to the typical enol form). DHF, thus formed, can react via aldol condensation with an additional molecule of glyoxylate (^{13}C labeling shown in red) to yield a six-carbon tricarboxylate intermediate **6***, which can (under high pH) rearrange to **7*** via cyanide mediated retro-aldol pathway. These intermediates **6***/**7*** can be attacked at the carbonyl by free hydroxide (under high pH) and undergo fragmentation to yield tartrates **8**/**8*** and oxalate **9**/**9***. This mechanism accounts for the primary pathway to tartrates, in which oxalate and tartrates are formed in a 1:1 ratio.

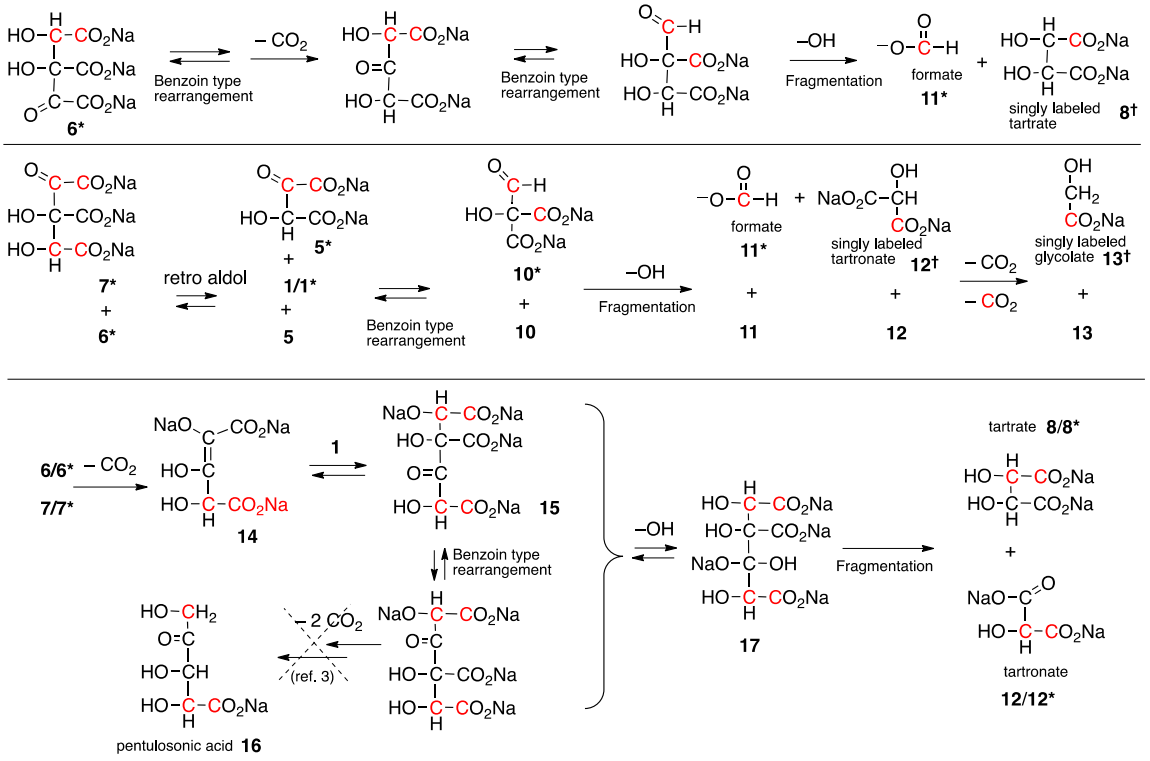


Figure 20 – Potential pathways to account for the formation of side products: tri-carboxylate intermediate **6(6*)/7(7*)** can undergo benzoin type rearrangements to a six-carbon aldehyde intermediate which fragments to yield singly labeled tartrates and formate. Alternatively **6(6*)/7(7*)** can undergo a retro-aldol generating DHF **5/5***, which under these high alkaline conditions, undergoes a benzoin type rearrangement to an aldehyde intermediate **10/10***. This intermediate can then react with hydroxide and fragment to form bicarbonate, formate **11/11***, tartronate **12/12*** and glycolate **13/13***. This second reaction pathway was identified from the experiments dealing with the heterogeneous decomposition of DHF alone to give formate and tartronate in 2M NaOH. This also accounts for the presence of formate in these samples; glycolate and oxalate are also produced by the hydroxide promoted fragmentation reaction of the keto form of DHF.

Alternatively **6(6*)/7(7*)** can undergo a retro-aldol reaction generating DHF **5/5***, which under these high alkaline conditions, undergoes a benzoin type rearrangement to an aldehyde intermediate **10/10***. This intermediate can then react with hydroxide and fragment to form bicarbonate, formate **11/11***, tartronate **12/12*** and glycolate **13/13***. This second reaction pathway was identified from the experiments dealing with the heterogeneous decomposition of DHF alone to give formate and tartronate in 2M NaOH. This also accounts for the presence of formate in these samples; glycolate and oxalate are

also produced by the hydroxide promoted fragmentation reaction of the keto form of DHF.

In addition, there is a side reaction that occurs when glyoxylate **1** is in a far higher concentration than DHF **5**, where tartrates **8** and tartronate **12** are formed in a 1:1 ratio. This side reaction (Scheme 4, bottom) begins at intermediate **6**/**7**, which decarboxylates to intermediate **14**, which then undergoes an aldol reaction with an additional molecule of glyoxylate **1** to form a seven carbon tricarboxylate **15**. Further benzoin type rearrangements followed by fragmentation gives rise to tartrates **8** and tartronate **12**.

The prevalence of this side reaction is well illustrated when considering the product suite of the heterogeneous glyoxoin reaction, versus the homogeneous case (Figure 11 A vs B). In the heterogeneous case carboxylate is the dominant signal and tartronate is produced in a ratio much closer to 1:1 with tartrate. In the homogeneous case, little carbonate is produced and oxalate and tartrate are in a near 1:1 ratio with little tartronate formation. An additional side reaction was identified where heterogeneous DHF decomposes to formate and tartronate in 2M NaOH which accounts for the presence of formate in these samples.

2.2 Kinetic and Thermodynamic Modeling of the Glyoxoin Reaction Pathway

To further enhance our understanding of this reaction pathway, specifically with the hopes of better understanding the difference between the low and high pH reaction pathways, we undertook to create a kinetic and thermodynamic model of this reaction pathway. Density functional theory thermodynamic modeling (DFT) of transition state and ground state energetics were employed at two levels. First we modeled each step in

our reaction pathway using DFT at the B3LYP/6-31* level in the gas phase. The purpose of this calculation was as an initial survey to validate that reasonable transition state energies existed in our proposed mechanism, prior to kinetic modeling. Second, we conducted reactions of 13C labeled glyoxylate and cyanide at various temperatures, which were monitored by quantitative 13C NMR for times up to four days, with spectra collected every 30 minutes. Integration of these spectra allowed the creation of kinetic profiles which were fitted to our reaction model, giving individual reaction constants. Arrhenius plots of these reaction constants allowed calculation of activation energies. Finally, in collaboration with the Leszczynski group at Jackson State University, we modeled both the deoxalation mechanism and lower pH decarboxylation mechanism using DFT at the M06-2X/6-311+G(d,p) level with PCM model water. The purpose of this final calculation was 1) to compare the kinetic results with a more realistic calculation, 2) to obtain a more granular view of the rate limiting step than afforded by the NMR kinetics, which by necessity encompassed several mechanistic steps, and 3) to understand the differences between the reaction at the two pHs,

2.2.1 Density Functional Theory Modeling (B3LYP/6-31G*)

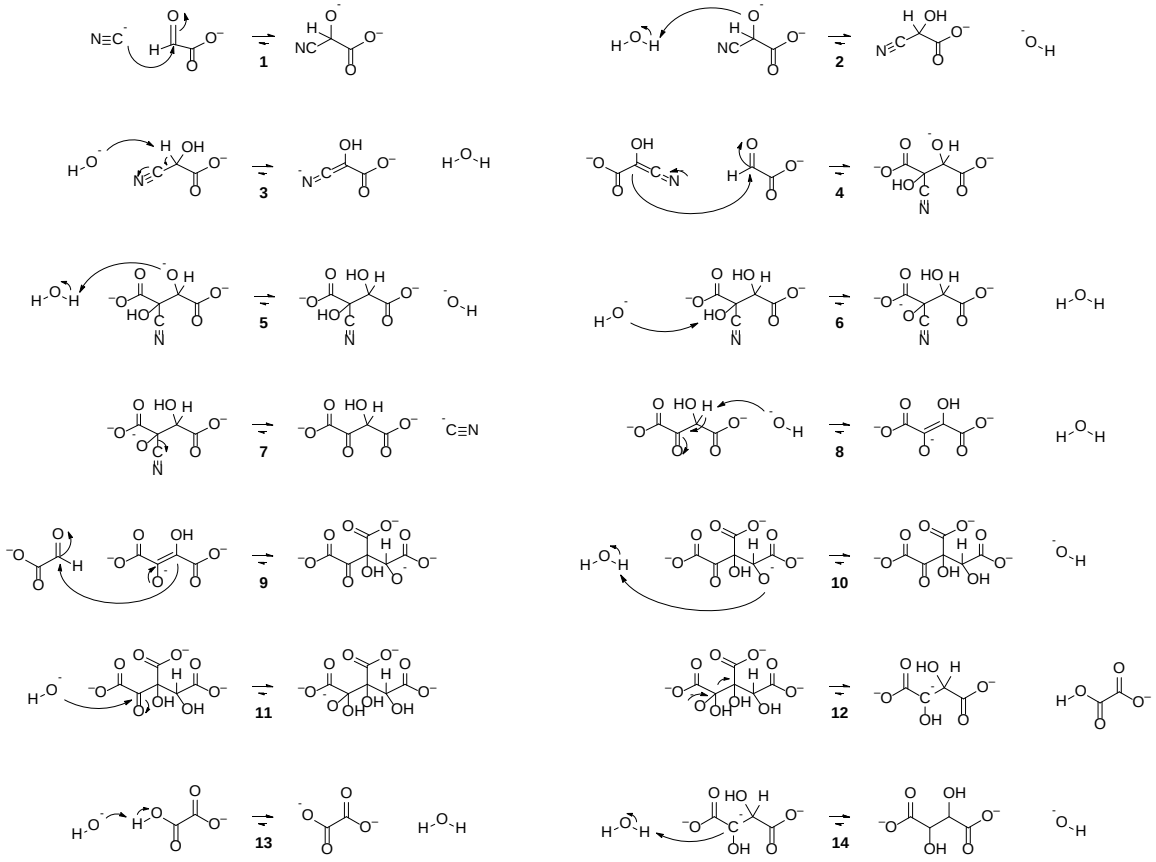


Figure 21 – Mechanism of Tartrate Formation: Individual Reactions for Modelling with Density Functional Theory Shown here are the 14 individual transitions considered in the transformation of glyoxylate and cyanide to tartrate. Combined these represent the simplest pathway from tartrate to glyoxylate as depicted in Figure 19.

To aid the process of modeling the reaction pathway, our proposed reaction mechanism described in Figure 19 was broken down to fourteen individual mechanistic steps shown in Figure 21. The reactants, products, and transition states of each of these reactions were modeled in our laboratory using density functional theory at the B3LYP/6-31G* level in the gas phase in the Spartan ’14 software package. The purpose of this modeling was to allow us to garner insight into the specific mechanistic steps taking place in this reaction, and to compare alternate possible mechanisms.

The results of this modeling are shown in Figure 23. The reactions shown in Figure 22 were also investigated, but rejected as steps in the main pathway due much higher energies as opposed to Reactions 6 and 7 from Figure 21.

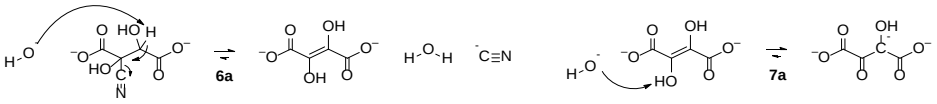
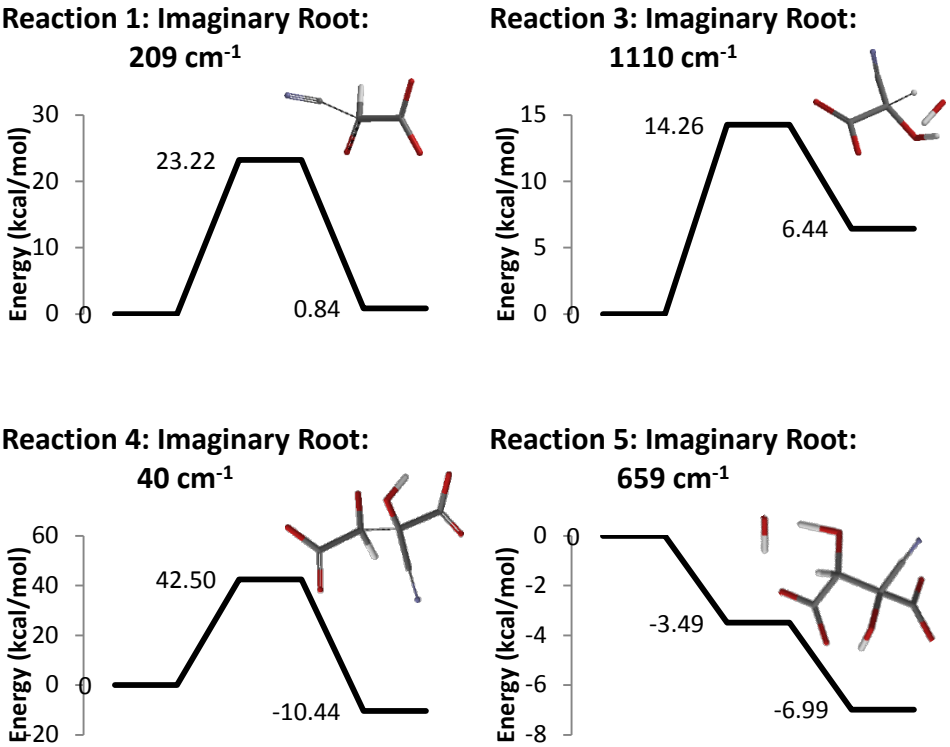
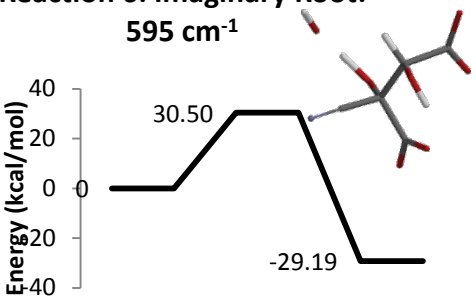


Figure 22 – Reaction steps rejected based on initial DFT screening.

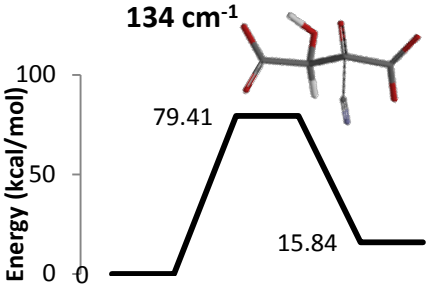
The higher reaction energy of Reaction 6a versus Reaction 6 is consistent with our inability to observe DHF as a reaction intermediate under any reaction conditions. Reaction 6 yields the keto form of dihydroxyfumarate rather than the standard -ene diol species. Deprotonation at the alpha carbon of this keto species will yield an activated intermediate capable of reacting with glyoxylate without ever forming the standard DHF isomer. Due to the simple basis set employed, further comparisons would be unfounded.



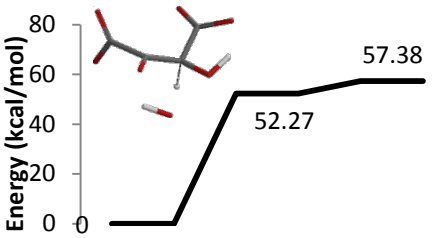
Reaction 6: Imaginary Root: 595 cm^{-1}



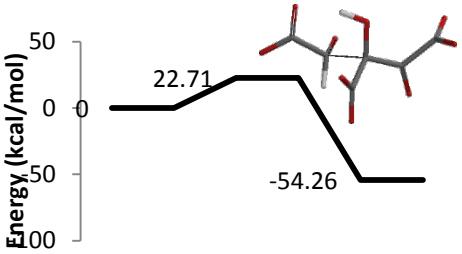
Reaction 7: Imaginary Root: 134 cm^{-1}



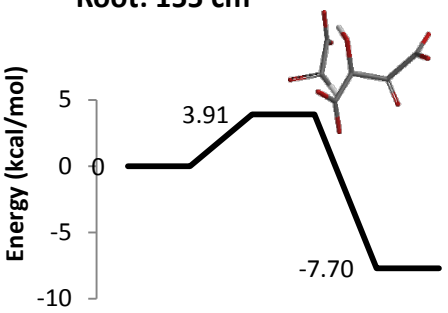
Reaction 8: Imaginary Root: 791 cm^{-1}



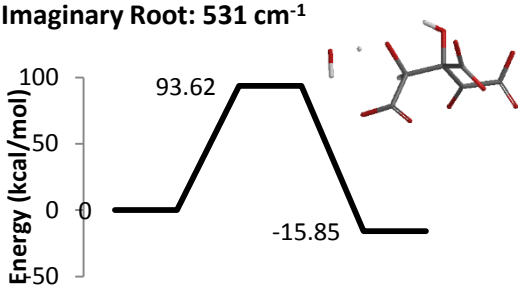
Reaction 9 - meso: Imaginary Root: 152 cm^{-1}



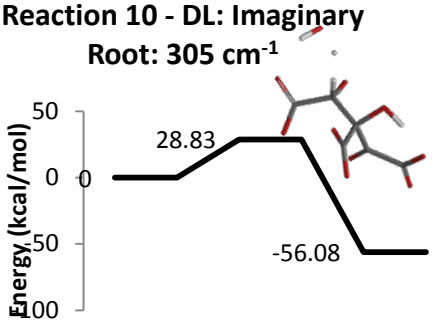
Reaction 9 - DL: Imaginary Root: 155 cm^{-1}



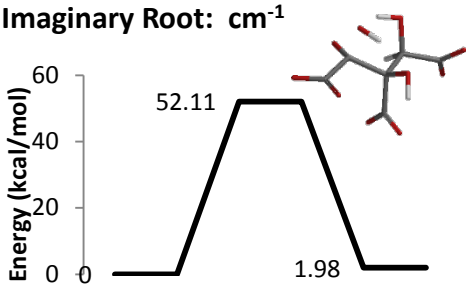
Reaction 10 - meso: Imaginary Root: 531 cm^{-1}



Reaction 10 - DL: Imaginary Root: 305 cm^{-1}



Reaction 11 - meso: Imaginary Root: cm^{-1}



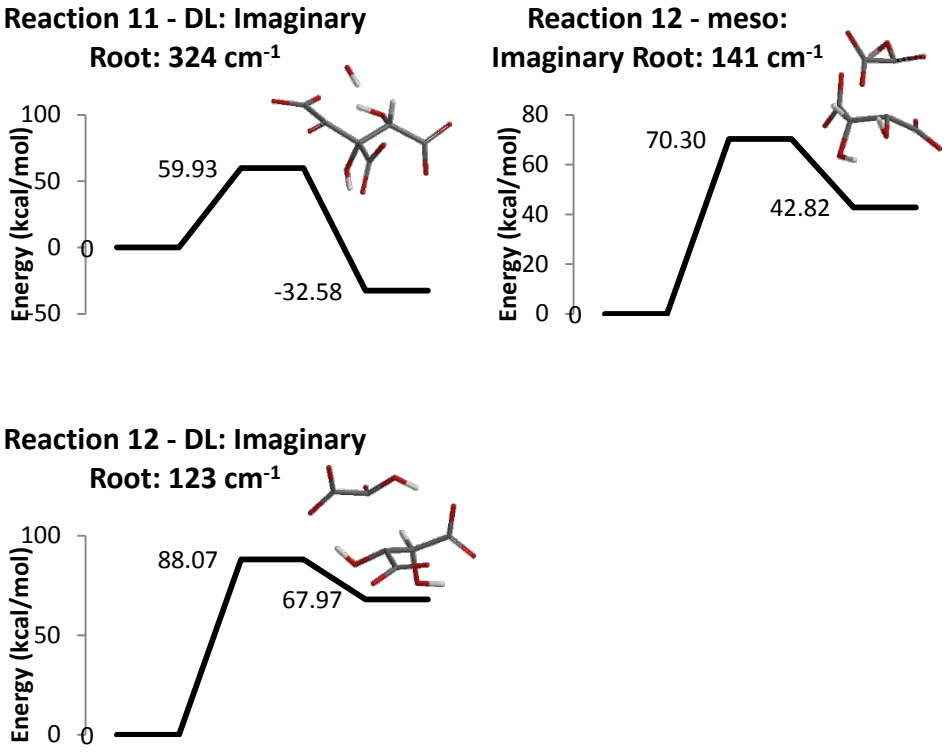


Figure 23 – Comparison of Results for the Gas Phase Calculation of the Energetics of the Glyoxoin Reaction.

2.2.2 Kinetic Modeling of Quantitative ^{13}C NMR Data

Having identified the reactions depicted in Figure 21, as the primary reactions responsible for production of tartrate from glyoxylate and cyanide, we undertook to create a kinetic model of the reaction. Data for this model was obtained by monitoring reactions of 0.1M ^{13}C glyoxylic acid, 0.01M ^{13}C NaCN, and 2.0M NaOH in situ using a quantitative ^{13}C NMR time course while conducting the reactions within the spectrometer. Quantitative ^{13}C spectra were collected every 25 minutes, allowing for tracking of all products and reactants, as well as the glyoxylate cyanohydrin intermediate (Figure 24). Data were collected at 25, 15, and 4 degrees centigrade. The collected NMR

signals were integrated, and reactant concentrations calculated from these integrals to create a kinetic profile. These data were then fitted using the DynaFit kinetic analysis software [11] to obtain numerical values for each of the pseudo rate constants for the reactions shown in Figure 25.

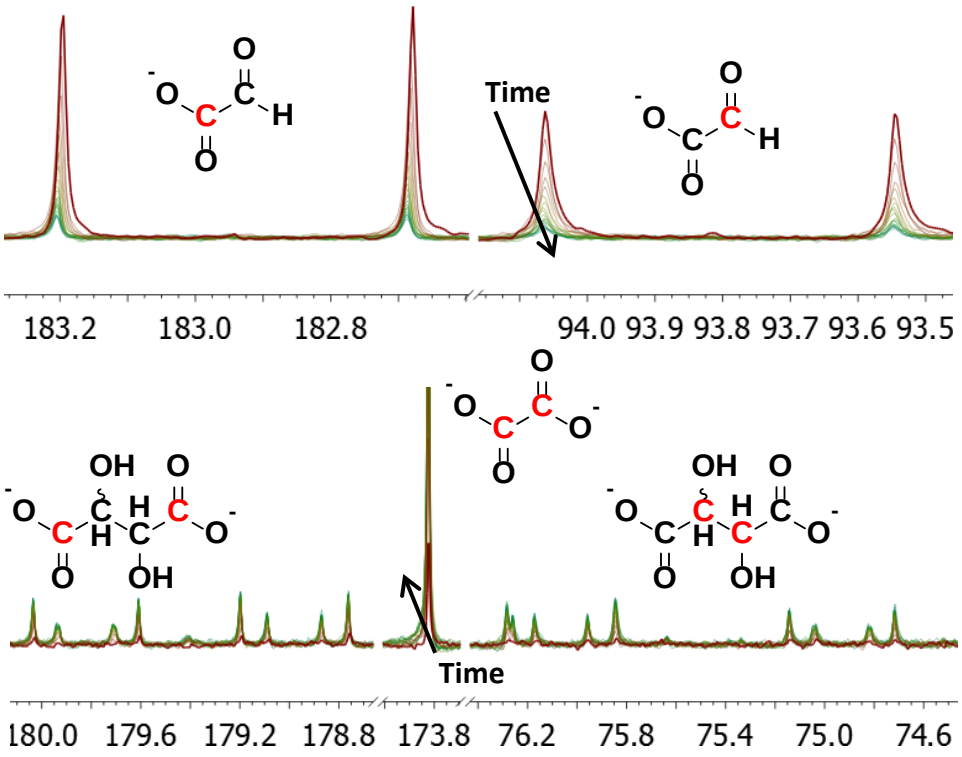


Figure 24 – Example NMR time course. Decrease of glyoxylate and increase of tartrates and oxalate are clearly visible as the reaction progresses. Not shown are peaks corresponding to cyanide, glyoxylate cyanohydrin, and sideproducts, all of which can be followed similarly.

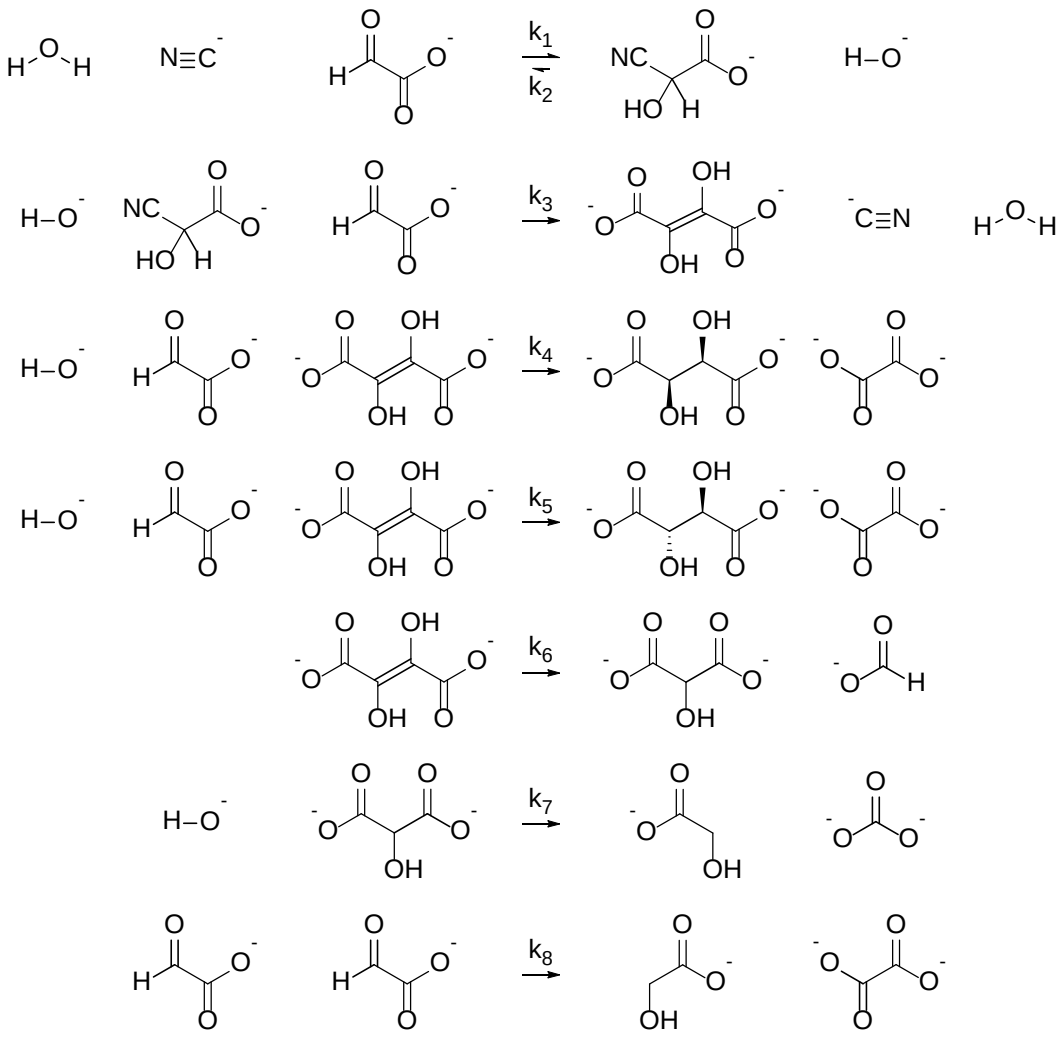


Figure 25 – Reactions and Psuedo Rate Constants Considered for Kinetic Modeling of Glyoxoin Reaction.

Fitted values were obtained from six reactions, two at each temperature of interest. The Arrhenius equation

$$\ln(k) = \frac{-E_a}{R} \cdot \frac{1}{T} + \ln(A)$$

was utilized to obtain activation energies for each reaction from a plot of the rate constants versus the inverse of temperature for each of the pseudo rate constants shown above[12, 13]. The Arrhenius plots for all reaction constants are shown in Figure 26. The activation energies and pre-exponential constants are tabulated in Table 3

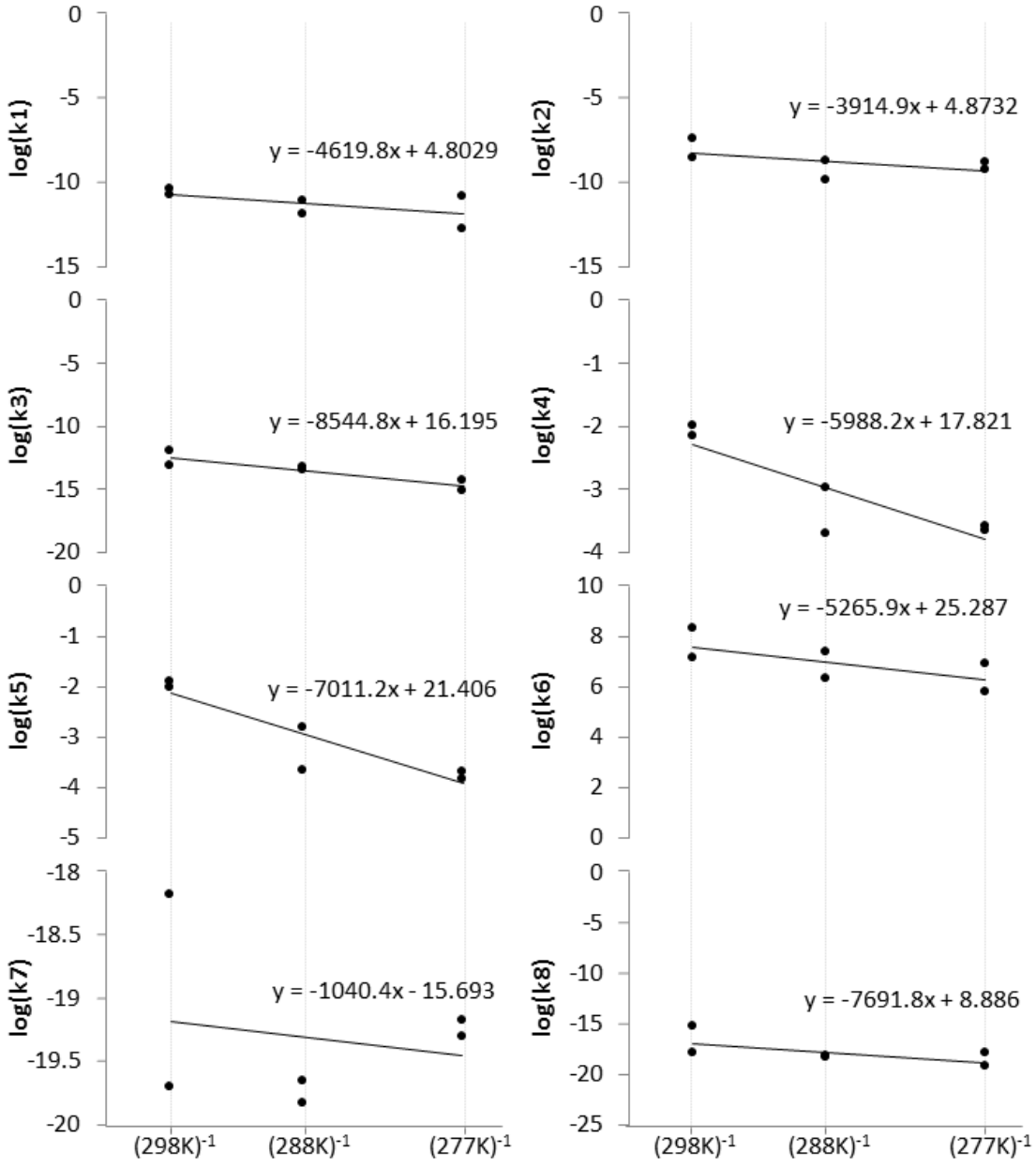


Figure 26 – Arrhenius Plots of Rate Constants for Reactions Leading to the Production of Tartrate and Side Products From Glyoxylate and Cyanide.

Table 3 – Activation Energies and Pre-exponential Constants of Glyoxoin Reaction Steps

	E_a (kcal/mol)	A (unit)
k1	9.179532	1.22E+02 $L^2 \cdot mol^{-2} \cdot s^{-1}$
k2	7.778934	1.31E+02 $L \cdot mol^{-1} \cdot s^{-1}$
k3	16.97861	1.08E+07 $L^2 \cdot mol^{-2} \cdot s^{-1}$
k4	11.89857	5.49E+07 $L^2 \cdot mol^{-2} \cdot s^{-1}$
k5	13.9313	1.98E+09 $L^2 \cdot mol^{-2} \cdot s^{-1}$
k6	10.46338	9.59E+10 s^{-1}
k7	2.06726	1.53E-07 $L \cdot mol^{-1} \cdot s^{-1}$
k8	15.2836	7.23E+03 $L \cdot mol^{-1} \cdot s^{-1}$

The Arrhenius data for this reaction show k_3, the rate of reaction of glyoxylate and the glyoxylate cyanohydrin to form dihydroxyfumarate is the rate limiting step of this reaction with an activation energy of 16.97 kcal/mol. This result is consistent with the rapid reaction of DHF and glyoxylate which, when reacted at concentrations of 0.1M or above, had progressed to completion in the time needed to conduct the NMR measurement. Were the rate controlling step after the reaction of DHF and glyoxylate, it would be expected that this reaction would proceed more slowly.

The calculated 15.28 kcal/mol for the Cannizzaro reaction of glyoxylate(k8) is in line with the 13.85 kcal/mol energy of formation of the formaldehyde Cannizzaro reaction reported in the system of 50/50 water/methanol[14]. The only rate constant of the eight fitted which does not conform to Arrhenius behavior is k7. The value for this constant is driven solely by the small amount of carbonate produced in these reactions. As shown in Figure 20, we have identified several routes to carbonate formation in these reactions. Only the decomposition of tartronate was considered here, based on the observation of carbonate and glycolate formation in DHF control reactions. The results of this kinetic analysis suggest that the carbonate is produced not by one dominant reaction, but rather a

suite of reactions contributing in similar amount. Unfortunately the low concentration of carbonate, especially coupled with the noise in the NMR measurements, makes deconvolving these reactions unfeasible.

2.2.3 Density Functional Theory Modeling (M06-2x/6-311+G(d,p))

While our kinetic measurements allowed us to identify the reaction of glyoxylate and its cyanohydrin as the rate limiting step in the glyoxoin reaction, the pseudo rate constant k3 actually encompasses five mechanistic steps (Figure 27). To identify the mechanistic rate limiting step, we conducted a higher level DFT calculation at the M06-2x/6-311+G(d,p) level with PCM water, for comparison to our kinetic results. These calculations were performed by our collaborators in the Leszczynski group at Jackson State University using the Gaussian ’09 computational package[15]. Alcohol proton transfers were omitted as they are expected to be facile. These results can be seen in Figure 28. From these calculations the overall ΔG of reaction is -67 kcal/mol, consistent with the observed boiling of high concentration reactions.

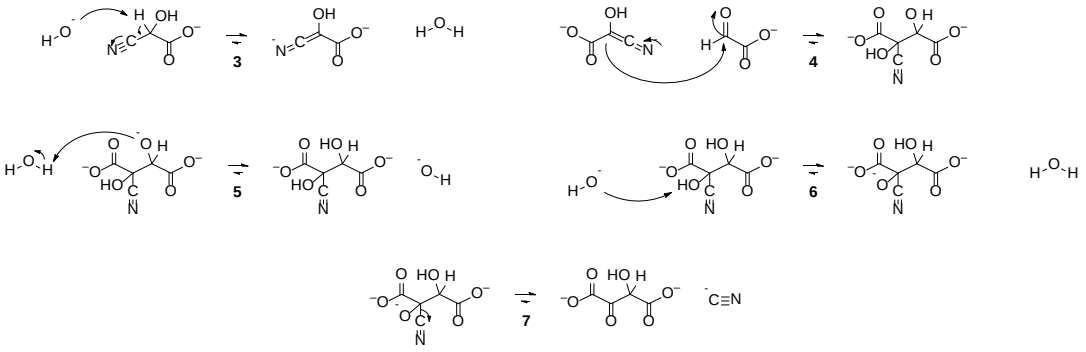
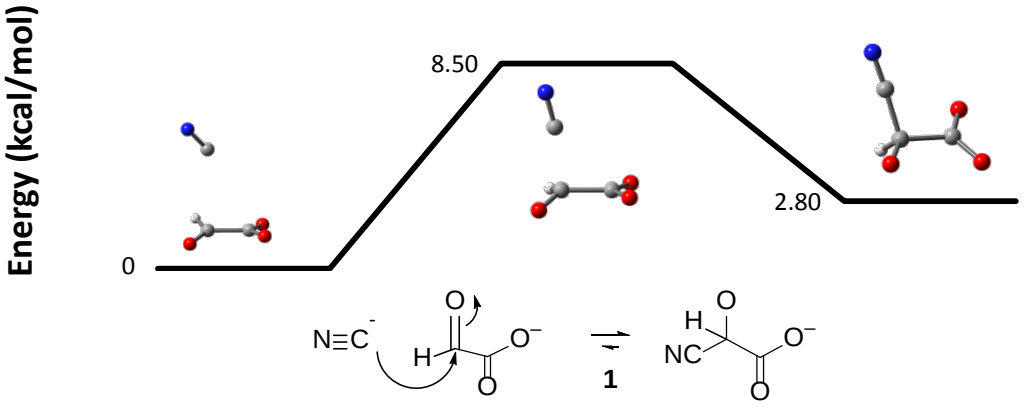
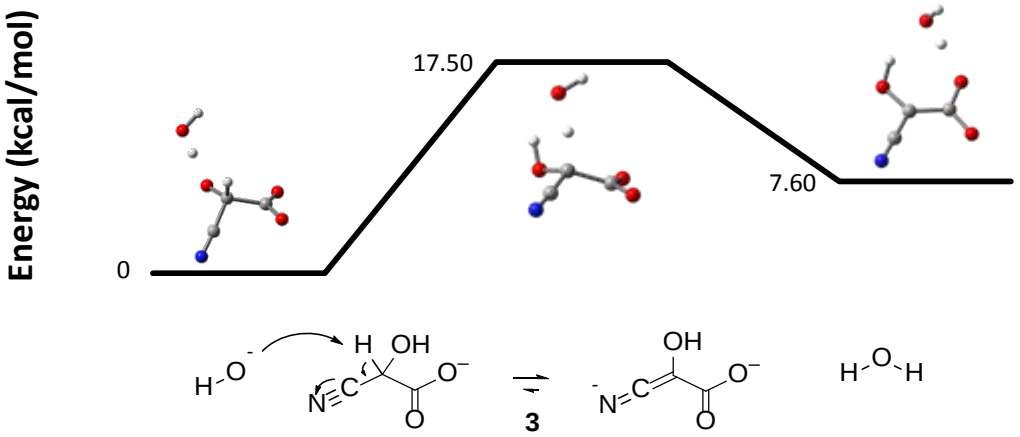


Figure 27 - Mechanistic Steps Encompassed by the Psuedo-Rate Constant k3

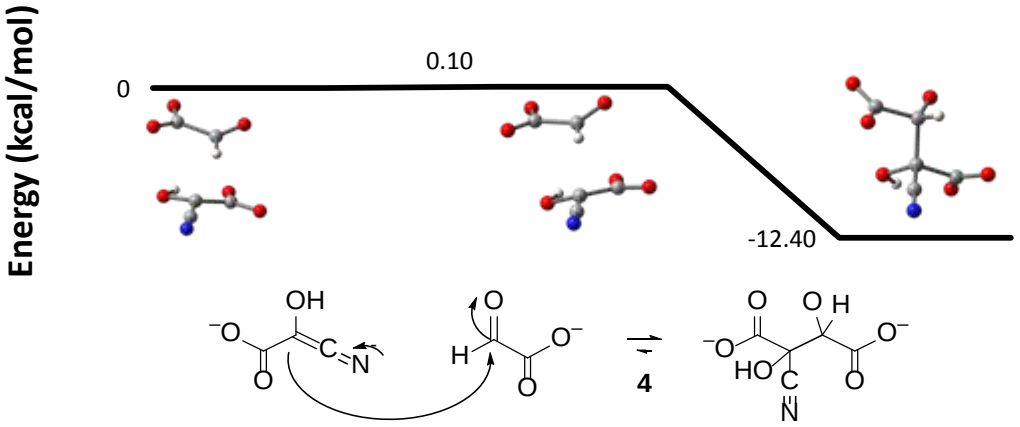
Reaction 1: Imaginary Root: 288 cm^{-1}



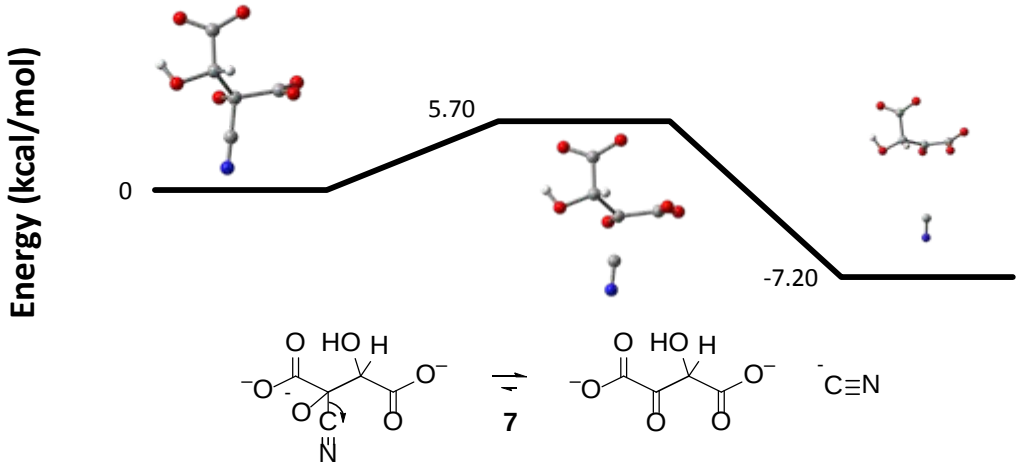
Reaction 3: Imaginary Root: 1250 cm^{-1}



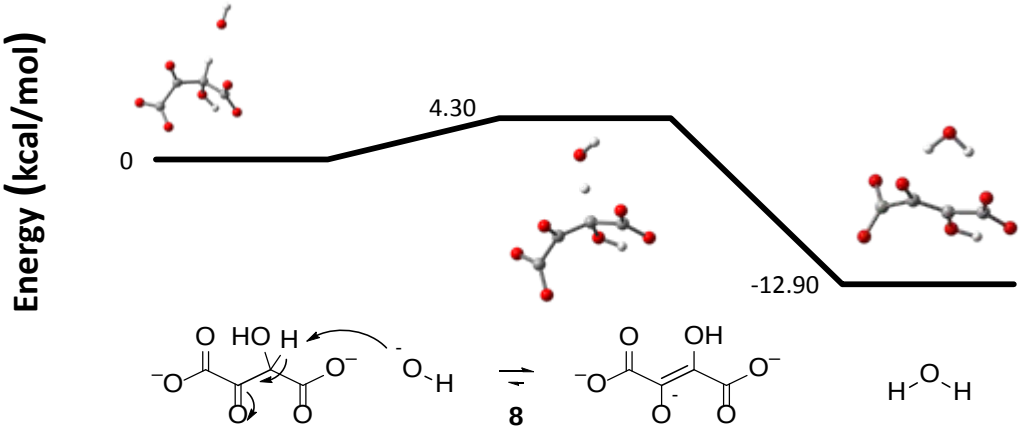
Reaction 4: Imaginary Root: 43 cm^{-1}



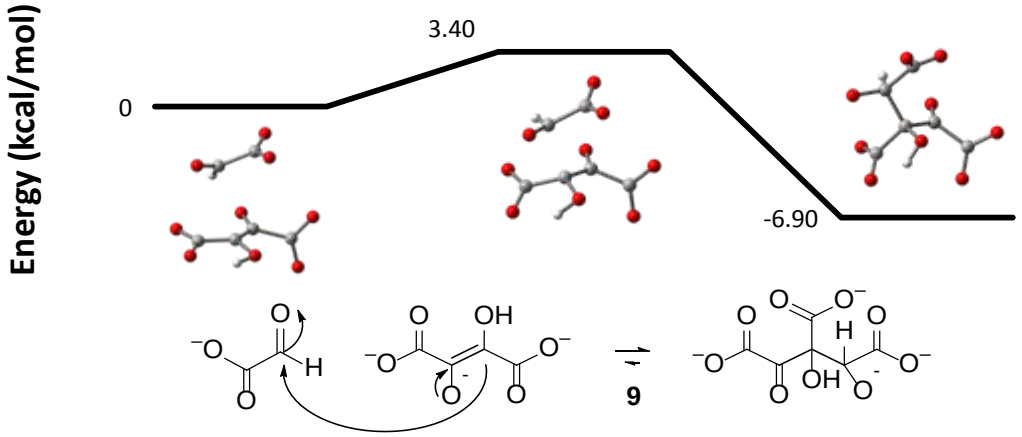
Reaction 7: Imaginary Root: 242 cm^{-1}



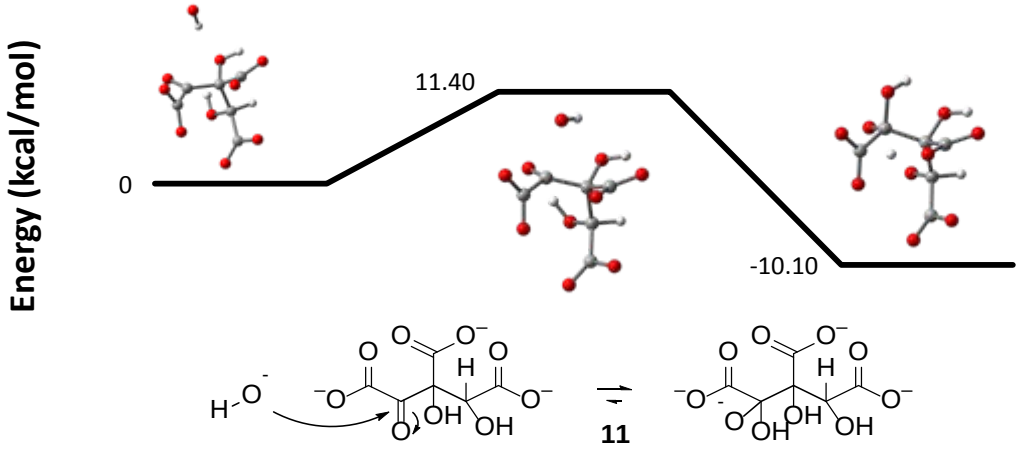
Reaction 8: Imaginary Root: 1263 cm^{-1}



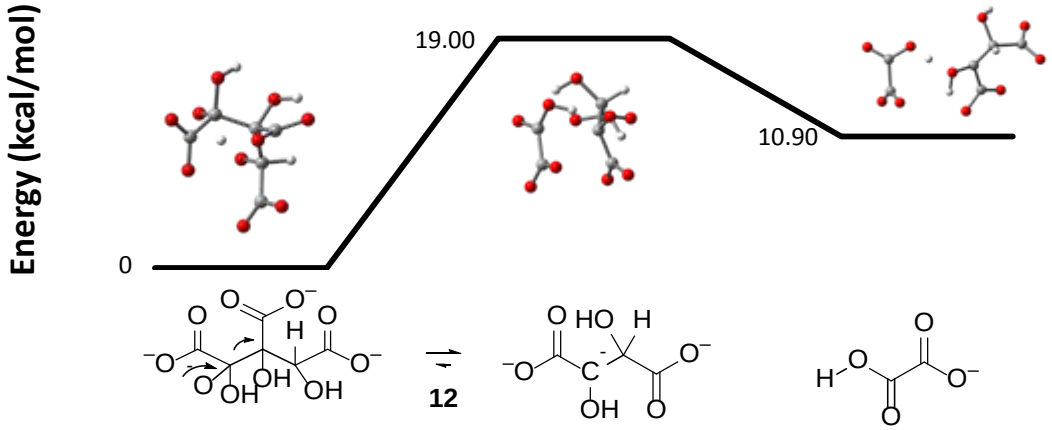
Reaction 9: Imaginary Root: 166 cm^{-1}



Reaction 11: Imaginary Root: 222 cm^{-1}



Reaction 12: Imaginary Root: 97 cm^{-1}



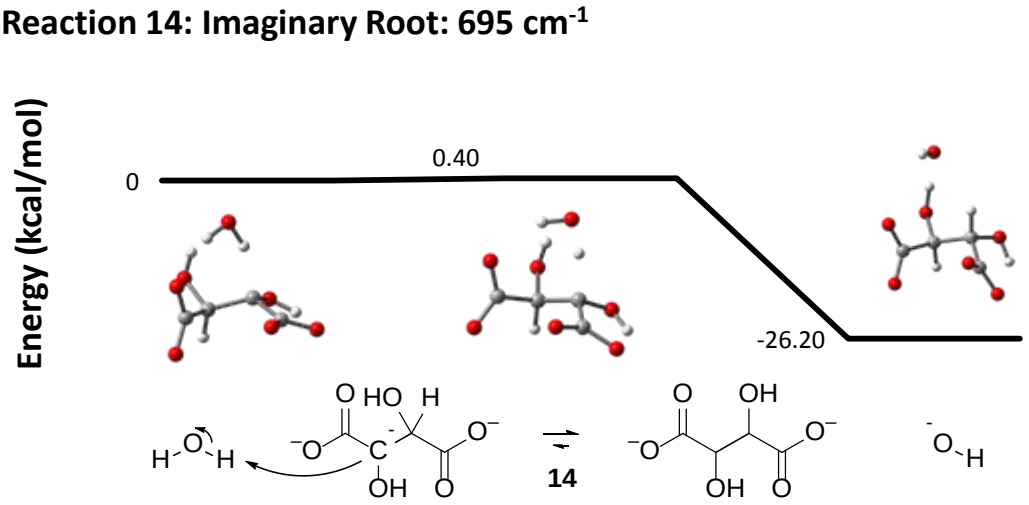


Figure 28 - Reaction Energetics and Geometries of the Glyoxoin Reaction Calculated Using DFT at the M06-2x/6-311+G(d,p) level with PCM water

These calculations interestingly suggest that mechanistic step 12 should be the rate limiting step for the glyoxoin reaction, with an energy barrier of 19 kcal/mol. However, the calculations also show that step 11 has an energy yield of 10.1 kcal, suggesting that molecules may not settle into the local minimum between steps 11 and 12. Also interesting is the suggestion that step 4 may have an extremely minimal energy barrier of 0.1 kcal/mol. This is an unlikely result, however, the imaginary transition state is along the reaction ordinate, and the carbon-carbon bond length of the forming bond at the transition is 2.61 angstroms, consistent with formation of a C-C bond, providing tangential proof that it may be accurate. The calculated energy barrier of 17.5 kcal/mol for step 3 is in very good agreement with the kinetic result of 16.97 kcal/mol. While the result for step 4 is somewhat suspect, the very good agreement between step 3 and the kinetic result, does suggest that the energy barrier of step 4 should be considerably lower than step 3.

Table 4 – Comparision of Kinetic and DFT Results

Kinetic Psuedo-step	Mechanistic Step	Activation Energy (kcal/mol) DFT	Kinetic Value
k1	1	**8.5**	**9.17**
	2	---	
k3	3	**17.5**	**16.97**
	4	0.1	
	5	---	
	6	---	
	7	5.7	
k4/k5	8	4.3	**11.89 / 13.93**
	9	3.4	
	10	---	
	11	**11.4**	
	12	19*	
	13	---	
	14	0.4	

The M06-2x/6-311+G(d,p) DFT results are in overall very good agreement with the kinetic results (Table 4). In addition to the good agreement between k3 and mechanistic step 3, the kinetic value for the calculated activation energy for the formation of the glyoxylate cyanohydrin at 8.5 kcal/mol is in very good agreement with the kinetic value of 9.17 kcal/mol. The agreement of the calculation with the calculated k4 and k5 is dependant on whether or not step 12 occurs as calculated, or if step 12 happens immediately after the excitation from the release of energy in step 11. If 11 is the rate limiting step with 12 happening immediately, then the results are in good agreement as well.

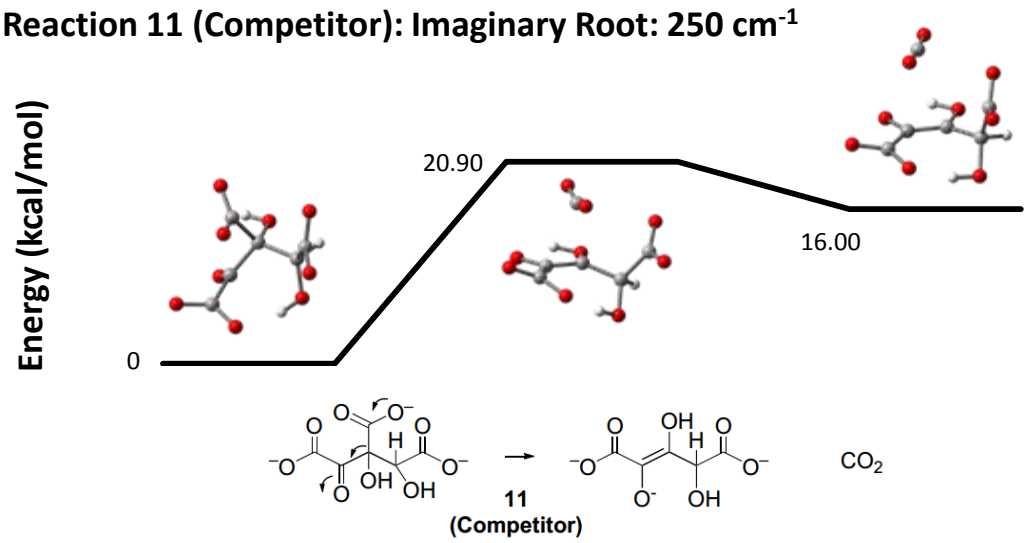


Figure 29 - Mechanistic Step 11, Competing Decarboxylation

In addition to the above calculation we also investigated the decarboxylation reaction shown by Sagi et al.[16] which competes with mechanistic step 11, shown in Figure 29. This decarboxylation has an activation energy of 20.9 kcal/mol as opposed to 11.4 kcal/mol of mechanistic step 11. This suggests that the "switch" between the decarboxylation and deoxalation reactions is governed solely by the concentration and nucleophilicity of hydroxide, rather than any catalytic effect thereof.

2.3 Summary

We have identified a plausibly prebiotic pathway for the formation tartrate from the reaction of glyoxylate and cyanide. All evidence indicates that this reaction proceeds by analogy to the benzoin reaction to produce dihydroxyfumarate. This dihydroxyfumarate then reacts with an additional equivalent of glyoxylate to form a six carbon tri-acid intermediate which fragments through a hydroxide promoted deoxalation pathway to form tartrate. We have thoroughly investigated the mechanisms by which this

reaction progresses through reactions of carbon labeled glyoxylate and cyanide with unlabelled DHF. We have shown that this cyanide mediated dimerization of glyoxylate at high pH produces tartrates with rapidly and consistently, even at low temperature, low concentration, and low solubility.

Kinetic and computational investigation of the reaction shows the rate limiting step to be the deprotonation of the glyoxylate cyanohydrin, underscoring the necessity of high pH for this reaction to proceed. At the pH required for this deprotonation, the concentration and nucleophilicity of hydroxide outcompetes the potential decarboxylation reaction of the six carbon triacid product of DHF and glyoxylate, instead proceeding through a deoxalation pathway.

The tartrate products of the glyoxoin reaction are stable under the high pH conditions required for the reaction to proceed, indicating potential as feedstock for further reactions. While it is true that high hydroxide concentrations are not often considered in conventional prebiotic line of thought, the widely investigated alkali hydrothermal vent system[17-19] is capable of obtaining pH as high as 12.6 which we have shown is sufficiently alkaline for this reaction to occur; however, this system is not without drawbacks[20]. Alternately, lakes fed by alkali springs can form reservoirs of naturally high hydroxide concentration [21], which could act as a natural reactor for this type of chemistry. Finally, another possible high alkaline environment exists in the interlayer framework of double-layered hydroxides (e.g. hydrotalcites), which have been shown to be especially conducive for uptake and concentration of anions (such as glyoxylate and cyanide) from dilute solutions.[22, 23]

Having a robust production of tartrates from glyoxylate may spur new avenues of research into the origins of biologically relevant small molecules. The known thermal dehydration of tartrates[24] results in oxaloacetate which is known to decarboxylate to give pyruvate[25], thus providing an entry into a prebiotic citric acid cycle. In fact, work by Cooper et al.[26] suggests that simply by attaining access to pyruvate, one gains access to all intermediates of the citric acid cycle with the exceptions of isocitrate and succinate. It is then reasonable to posit that tartrates could have acted as a source of small molecules, which could become part of an emerging proto-metabolic process (e.g. reductive citric acid cycle [27, 28], Figure 30). Thus, the glyoxoin reaction (and the "glyoxylate scenario") could serve as a plausible alternative start for rudimentary chemical evolution.

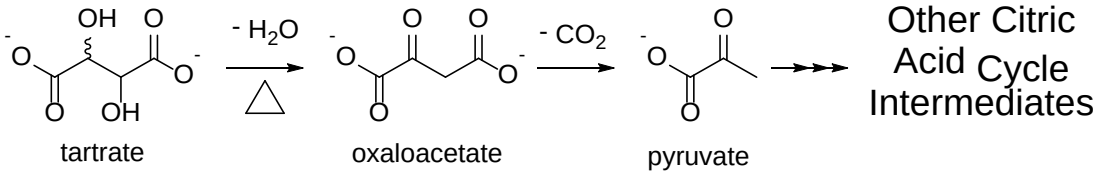


Figure 30 - Simple Conversions Leading from Tartrate to Intermediates of the Citric Acid Cycle

2.4 Experimental

Materials: Lithium hydroxide, sodium cyanide, glyoxylic acid monohydrate, and dihydroxyfumaric acid dihydrate were all purchased from Sigma Aldrich. Sodium hydroxide was obtained from EMD Millipore. Di-labeled ^{13}C glyoxylic acid (CLM-6027-PK; Lot PR-11812) and singly labeled ^{13}C sodium cyanide (CLM-378-1; Lot PR-20741) both with 99% isotopic substitution were purchased from Cambridge Isotope Laboratories Inc.

NMR: ^{13}C NMR data were obtained using either a Bruker Avance II-500 spectrometer or a Bruker Avance III-400 spectrometer. All NMR data were obtained using a 5mm probe, with a 1mm d_4-MeOH capillary as external standard and lock solvent. Data were collected using an inverse gated, 30 degree, ^{13}C pulse sequence with a D1 of 45 seconds.

Quantitative NMR results were assured mathematically, using the methods given by Traficante [29]. For this analysis, product T_1 measurements were obtained from an inversion recovery experiment. The longest T_1's were those of oxalate and formate at 32 seconds. Based on this T_1 and a 30 degree pulse, we calculate a 95% relaxation of these species with a delay time of 45 seconds. At 300 seconds this relaxation is 99.998%. This indicates that an error of no larger than 5% will be observed with a delay time of 45 seconds. This mathematical result was also tested experimentally as shown in Figure 31. These NMR conditions allow us to collect reasonable quantitative data in 30% of the time required under the conditions of a 90 degree pulse and $5T_1$ relaxation time. Quantitative time courses were obtained by using this pulse sequence repeatedly, while the NMR internal temperature was controlled. For each time point (~25 minute spacing) 64 scans were collected at a delay of 45 seconds, the data saved, and a new experiment begun, allowing the creation of kinetic reaction profiles.

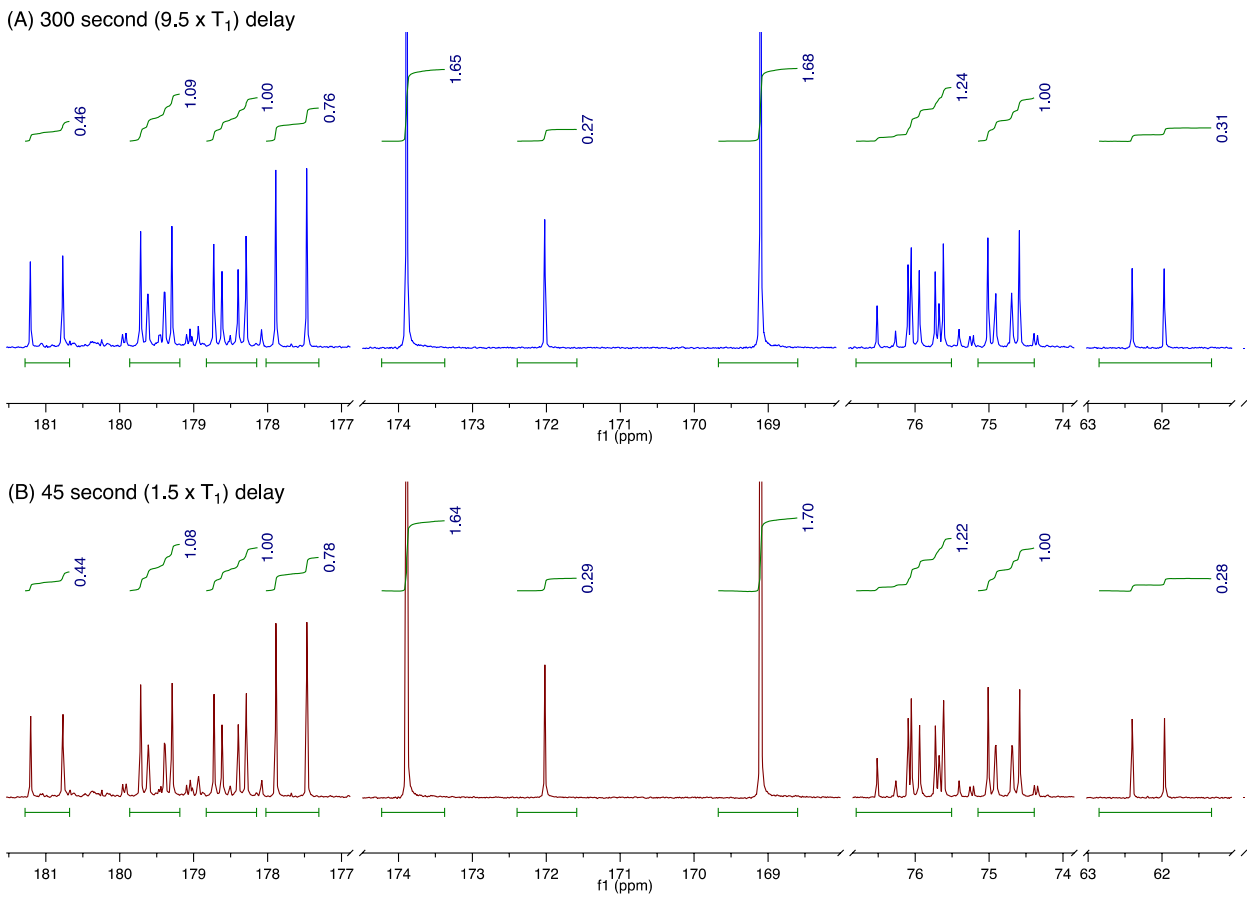


Figure 31 - Comparison of quantitative 30 and 90 degree NMR pulses. ^{13}C NMR spectra of the reaction of 1 M ^{13}C Glyoxylate with 0.01 M ^{13}C NaCN after two weeks. Both NMR spectra were obtained using an inverse gated pulse sequence with a relaxation delay. In (A) a relaxation delay of 300 seconds or 9.5 times the largest T_1 was used. In (B) the delay was 45 seconds. Inspection of the integration values shows a very good agreement between the two data sets, indicating the sufficiency of a 45 second relaxation time under these conditions. See Methods section for more information.

pH: pH was measured using Aldrich Whatman brand PANPEHA pH Paper (Sensitivity of 0.5 pH unit in the 0 to 9 range; 1 pH unit in the 9 to 14 range). At high pH, measurements using glass electrodes become inaccurate; therefore, pH paper was used.

Mass Spectrometry: Mass Spectrometry data were collected using a Waters 3100 Mass Detector without chromatographic separation. The electrospray ionization injector was operated in both positive and negative mode. The mobile phase was a 1:1 water/MeOH mixture with 1mM ammonium formate.

Stock Solutions: Stock solutions of LiOH and NaOH were prepared by weighing the required mass of the hydroxide into a browned glass bottle. The necessary amount of LCMS grade H_2O was added as measured by graduated cylinder. The bottle was then closed and subjected to sonic mixing until complete dissolution was achieved. Neither degassing nor inert atmosphere was required.

Reaction of Glyoxylate with Cyanide: Glyoxylic acid monohydrate (0.1 to 10 mmol, 10 to 100 equivalents) and sodium cyanide (0.01 to 1mmol, 1 equivalent) were each weighed in individual 6 dram vials. 10 mL of 2.0 M NaOH or 2.0 M LiOH were then added to the vial containing the sodium cyanide. This vial was mixed until no solid remained. The solution was then transferred by syringe to the vial containing the glyoxylic acid. The cyanide solution was added down the side of the vial to the glyoxylic acid over the course of approximately 20 seconds, then closed and the vial was cooled (by running cool water over the vial) as necessary to avoid with excess heating. The pH measurements of the resulting solutions were ≈ 14.

Heterogeneous Reaction of Dihydroxyfumarate with Glyoxylate and Cyanide: Dihydroxyfumaric acid dihydrate (0.1 to 5 mmol, 5 equivalents), glyoxylic acid monohydrate (0.1 to 5 mmol, 5 equivalents) and sodium cyanide (0.01 to 1mmol, 1 equivalent) were each weighed in individual 6 dram vials. 5 mL of 2.0 M NaOH were then added to the vial containing the sodium cyanide. This vial was mixed until no solid remained. In a separate vial an additional 5 mL of 2.0 M aq. NaOH were then added to the vial containing the glyoxylic acid. The glyoxylate and sodium cyanide solutions were then added to the dihydroxyfumaric acid. The glyoxylate solution was added down the

side of the vial to the dihydroxyfumaric acid over the course of approximately 20 seconds, then closed and the vial was cooled (by running cool water over the vial) as necessary to avoid with excess heating.. The pH of the resulting solution was ≈ 14.

Homogeneous Reaction of Dihydroxyfumarate with Glyoxylate: Dihydroxyfumaric acid dihydrate (0.1 to 5 mmol, 1 equivalent) and glyoxylic acid monohydrate (.1 to 5 mmol, 1 equivalent) were each weighed in individual 6 dram vials. 5 mL of 2M NaOH were then added to the vial containing the glyoxylic acid monohydrate. This vial was mixed until no solid remained. In a separate vial, 5 mL of 2.0 M aq. LiOH were then added to the vial containing the dihydroxyfumaric acid. This vial was mixed until no solid remained. The glyoxylate solution was then transferred into the vial containing the DHF solution, the vial closed, and shaken. Alternately, adding DHF to glyoxylate resulted in no noticeable differences. The pH of the resulting solutions were ≈ 14 by pH paper.

Density Functional Calculations Using Spartan: A molecular model was constructed of each intermediate and transition state as depicted in Figure 21. The energetics of the transition states were dominated by coulombic repulsion of the negatively charged species, and as such, the initial energetics simply represented the number of negatively charged moieties present in the model. To combat this, energies of solvation were calculated for each intermediate and incorporated into the reaction energy profiles. Energies of solvation for the transition states could not be calculated, and instead were substituted with a simple average of the energy of solvation of reactants and products. While this is not the most rigorous approach, neither is the rudimentary basis set

employed, and this method was adequate to correct for the coulombic repulsion and allow comparison of transition states without charge repulsion overwhelming other factors.

Fitting of kinetic data using DynaFit: Kinetic data were obtained from the integrals of NMR reaction time courses. These integrals were normalized to the number of carbons represented by each, and then used to obtain a percent abundance of each species in the reaction. These percent abundances were then converted to concentrations based on the known content of ^{13}C carbon in each reaction vessel. These concentrations were then prepared and entered into DynaFit which uses a nonlinear regression technique to fit the data to the chemical model. A Monte-Carlo analysis was also utilized in DynaFit, where each reaction time course was fit 1000 times while initial guesses for the values of each reaction rate were varied. The average value after removal of the 5% furthest outliers from these Monte-Carlo runs is presented herein.

2.5 References

[1] SALADINO R, et al. Photochemical synthesis of citric acid cycle intermediates based on titanium dioxide. Astrobiology 2011;11:815-24.

[2] ESCHENMOSER A. On a hypothetical generational relationship between HCN and constituents of the reductive citric acid cycle. Chem Biodivers 2007;4:554-73.

[3] LAPWORTH A. XCVI.—Reactions involving the addition of hydrogen cyanide to carbon compounds. Journal of the Chemical Society, Transactions 1903;83:995-1005.

[4] AUTHOR. Selectively separating oxalic, tartaric, glyoxylic and erythronic acids from aqueous solutions containing the same. Google Patents; 1976.

[5] CANNIZZARO S. Ueber den der Benzoësäure entsprechenden Alkohol. Liebigs Annalen 1853;88:129-30.

[6] GEISSMAN TA. The Cannizzaro Reaction. Organic Reactions. 2 ed: John Wiley & Sons, Inc.; 1944. p. 94-113.

[7] SOCHA R, et al. Homogeneously catalyzed condensation of formaldehyde to carbohydrates: VII. An overall formose reaction model. Journal of Catalysis 1981;67:207-17.

[8] ESCHENMOSER A. The search for the chemistry of life's origin. Tetrahedron 2007;63:12821-43.

[9] ARMSTRONG FB, et al. Stereochemistry of the reductoisomerase and αβ-dihydroxyacid dehydratase-catalysed steps in valine and isoleucine biosynthesis. Observation of a novel tertiary ketol rearrangement. J Chem Soc, Chem Commun 1974:351-2.

[10] CROUT DHG, HEDGECOCK CJR. The base-catalysed rearrangement of α-acetolactate (2-hydroxy-2-methyl-3-oxobutanoate): a novel carboxylate ion migration in a tertiary ketol rearrangement. J Chem Soc, Perkin Trans 1 1979:1982-9.

[11] KUZMIČ P. DynaFit—a software package for enzymology. Methods in enzymology 2009;467:247-80.

[12] ARRHENIUS S. Über die Reaktionsgeschwindigkeit bei der Inversion von Rohrzucker durch Säuren. Zeitschrift für physikalische Chemie 1889;4:226-48.

[13] ARRHENIUS S. Über die Dissociationswärme und den Einfluss der Temperatur auf den Dissociationsgrad der Elektrolyte: Wilhelm Engelmann; 1889.

[14] SWAIN CG, et al. Mechanism of the Cannizzaro reaction. J Am Chem Soc 1979;101:3576-83.

[15] AUTHOR. Gaussian 09. Wallingford, CT, USA: Gaussian, Inc.; 2009.

[16] SAGI VN, et al. Exploratory Experiments on the Chemistry of the “Glyoxylate Scenario”: Formation of Ketosugars from Dihydroxyfumarate. J Am Chem Soc 2012;134:3577-89.

[17] MARTIN W, RUSSELL MJ. On the origin of biochemistry at an alkaline hydrothermal vent. Phil Trans R Soc B 2007;362:1887-926.

[18] MARTIN W, et al. Hydrothermal vents and the origin of life. Nature Rev Microbiol 2008;6:805-14.

[19] GLANSDORFF N, et al. The origin of life and the last universal common ancestor: do we need a change of perspective? Res Microbiol 2009;160:522-8.

[20] MILLER SL, BADA JL. Submarine hot springs and the origin of life. 1988.

[21] PEDERSEN K, et al. Distribution, diversity and activity of microorganisms in the hyper-alkaline spring waters of Maqarin in Jordan. Extremophiles 2004;8:151-64.

[22] BOCLAIR J, et al. Cyanide Self-Addition, Controlled Adsorption, and Other Processes at Layered Double Hydroxides. Orig Life Evol Biosph 2001;31:53-69.

[23] RIVES V. Layered double hydroxides: present and future: Nova Publishers; 2001.

[24] DANIELÁCHATTAWAY F, DAVIDÁPARKES G. LXXIX.—The formation of derivatives of oxalacetic acid from tartaric acid. Journal of the Chemical Society, Transactions 1923;123:663-9.

[25] WILEY RH, KIM KS. Bimolecular decarboxylative self-condensation of oxaloacetic acid to citrolyformic acid and its conversion by oxidative decarboxylation to citric acid. The Journal of Organic Chemistry 1973;38:3582-5.

[26] COOPER G, et al. Detection and formation scenario of citric acid, pyruvic acid, and other possible metabolism precursors in carbonaceous meteorites. Proceedings of the National Academy of Sciences 2011;108:14015-20.

[27] MOROWITZ HJ, et al. The origin of intermediary metabolism. Proceedings of the National Academy of Sciences 2000;97:7704-8.

[28] SMITH E, MOROWITZ HJ. Universality in intermediary metabolism. Proceedings of the National Academy of Sciences of the United States of America 2004;101:13168-73.

[29] TRAFICANTE DD. Optimum tip angle and relaxation delay for quantitative analysis. Concepts Magn Reson 1992;4:153-60.

CHAPTER 3:
PRODUCTION OF SUGARS AND SUGAR ACIDS

Linear sugars are an important class of biomolecules in modern biology. From the phosphoribose and deoxyribose chains that make up the backbone of our genetic material, to the massive cellulose lattices that comprise the trunks of the largest life forms on the planet, linear sugars provide the structural basis for life. If one assumes that the earliest informational polymers had this same reliance on linear sugars for their structural backbone, the question of how such a large and selective quantity of linear sugars was produced abiotically becomes one of extreme interest.

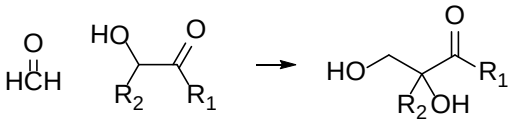
As discussed in Chapter 1, the formose reaction[1] has long been considered the gold standard in prebiotic production of sugar molecules. This is due both to the vast number of carbohydrates produced in the reaction and also to the assumed abundant atmospheric production of the starting material, formaldehyde, irrespective of whether the Hadean Earth had a reduced or oxidized atmosphere. [2-4]. While this reaction does demonstrate formation of a staggering number of sugars, there is no particular selectivity toward any given product[5]. In fact, the major downside of this reaction is the propensity for formation of branched sugars, which are not structurally useful[6], and therefore lack a major utility for emerging biochemistry. While many efforts have been made to direct the formose reaction to greater selectively for linear sugars, branched sugar production is reduced, but not eliminated[7]. Another impediment, of perhaps greater import, is the sugar products of the formose reaction, are not stable under the high pH conditions necessary for the reaction to proceed[8].

In this chapter, I will present my work on two alternatives to the formose reaction based on the chemistry of glyoxylic acid. These two reactions occur under different

environmental conditions and result in different suites of products each with their own advantages and drawbacks.

The first of these reactions (the glyoxylose reaction) is a direct analogy to the formose reaction, only with the formose feedstock replaced by glyoxylate (Figure 32). Both the glyoxylose and the formose reaction effect the same fundamental transformation of carbohydrates, the addition of a hydroxymethyl functionality; and both are dependant on a catalytic amount of initial carbohydrate for the reaction to progress. However, there are important differences in the mechanism arising from the carboxylic acid group in glyoxylate. The presence of this group makes glyoxylate a bulkier electrophile, decreasing the likelihood of reactions leading to branched products. Once a glyoxylate adduct is formed it must undergo carbonyl migration (a.k.a. a Lobry de Bruyn – van Ekenstein transformation[9]) followed by a decarboxylation. This step provides a second check against branched product formation; if a glyoxylate is added to form a branched product, it is impossible for the carbonyl to migrate to the β position preventing decarboxylation and further reaction. These two effects combine to make the glyoxylose reaction proceed more favorable toward linears sugars as compared to the formose reaction. Further benefit is gained from the fact that the glyoxylose reaction proceeds at pHs much closer to neutral (<9), an environment more favorable to the stability of the product sugars[10]. The mechanism and prebiotic implications of the glyoxylose reaction are discussed in greater depth in section 3.1.

Formose Pathway



Glyoxylose Pathway

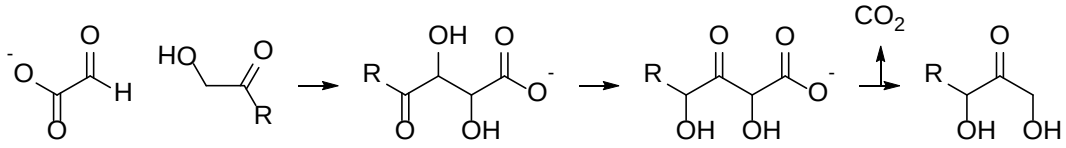


Figure 32 - Comparison of Formose and Glyoxylose Pathways for Sugar Elongation

The second reaction discussed in this chapter is the high pH reaction of dihydroxyfumarate with aldoses to form sugar acids. This reaction utilizes the same deoxalation chemistry identified in Chapter 2 as the mechanism by which DHF and glyoxylate react to form tartrates. Intrigued by the difference between the tartrate producing reaction of DHF and glyoxylate at high pH shown in Chapter 2, and the dihydroxyacetone producing reaction at pH ~ 8 reported by Sagi et al., we set out to investigate high pH reactions of DHF and the other aldehydes investigated by Sagi. We found that DHF reacts with paraformaldehyde to form glyceric acid, and with glycolaldehyde to form tetronic acids (Figure 33). This trend of chain elongation and oxidation continued even with aldoses up to glucose, however the reactivity did decrease as chain length increased, likely due to the reduced electrophilicity of the cyclic forms of longer sugars. Prebiotically, this reaction provides a means of chain growth and protection for small aldehydes; the carboxylate products of this reaction are much more stable under alkaline conditions than the starting materials. This could be employed in a manner analogous to the Kiliani-Fischer synthesis[11] in a cycle of additions to form sugar acids, and reductions to aldose, although prebiotically plausible conditions for

reduction of the carboxylate to an aldehyde would have to be identified. Alternately, the aldaric acid could be oxidized and undergo a decarboxylation to return to a ketose. Montmorillonite has been shown to effect a similar oxidative decarboxylation of isocitrate and may be a profitable avenue of research[12]. These high pH reactions of DHF are discussed in depth in Section 3.2.

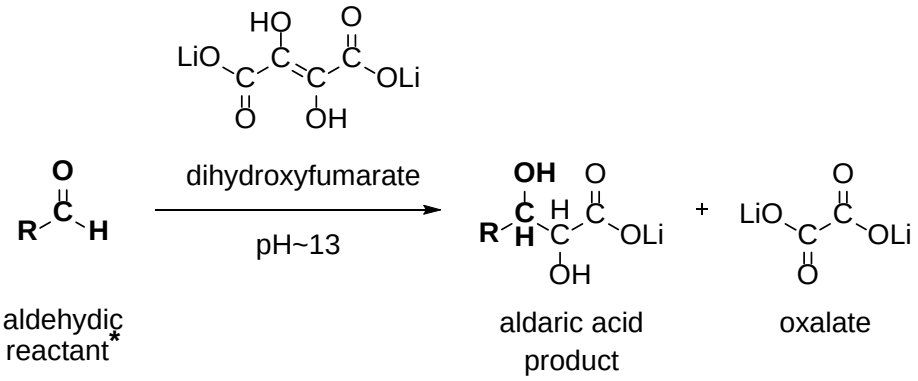


Figure 33 – Simplified Schematic of Reaction of Dihydroxyfumarate with Aldehydes

3.1 The Glyoxylose Reaction

As an alternative to the formose reaction we propose that addition of glyoxylate to ketoses can lead to formation of the next higher ketose through a pathway that partially mimics modern natural ketose formation (Figure 34). In modern biology ribulose is formed from an oxidized six carbon species, originating from glucose[13]. In our "glyoxylose" pathway this same oxidized six carbon species can be produced from an aldol reaction of erythrulose and glyoxylate, followed by a carbonyl migration. In fact, this glyoxylose reaction should, in principle, be applicable to the stepwise chain growth of any ketose. In this section I will demonstrate chain growth of ketoses from 3 to 6 carbons as well as application of this reaction to glycoloaldehyde.

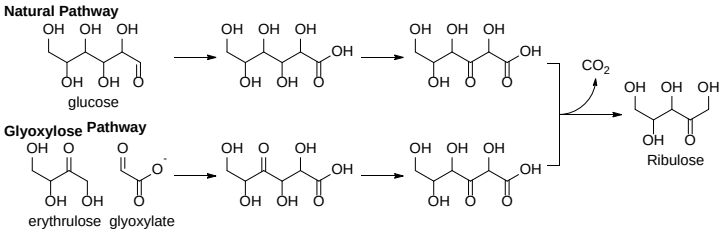


Figure 34 - Comparison of Glyoxylose and Natural Pathway of Ribulose Formation

3.1.1 Reaction of Glyoxylate with Dihydroxyacetone

We began our investigations by exploring the reaction of glyoxylate **3** with the simplest ketose, dihydroxyacetone (DHA) **5**. Initial investigations of the reaction were conducted by combining 0.5 M of each reactant in LCMS grade water and adjusting the initial pH to near 8.5 by addition of NaOH. After 1 day, the pH of the reaction had decreased to ca. 4.5. By ^{13}C NMR some conversion was observed, however glyoxylate **3** and DHA **5** persisted as major signals. The reaction pH was readjusted to 8.5 daily by addition of NaOH and monitored by ^{13}C NMR. After 7 days (Figure 36), the reaction was found to form primarily 4-oxo-pentulosonic acid **6**. While **6** was the expected product, the acidification of the reaction mixture is not explained by the 4-oxo-pentulosonic acid **6** producing reaction. Also observed in this reaction were peaks consistent with addition of a second glyoxylate **3** to 4-oxo-pentulosonic acid **6** to form 4-oxo-heptulosonic diacid **7**, shown as a side reaction in Figure 35. Not observed in this NMR were peaks corresponding to tetrulose **10** or carbonate, either of which would indicate that carbonyl migration of 4-oxo-pentulosonic acid **6** to 3-oxo-pentulosonic acid and subsequent decarboxylation was occurring.

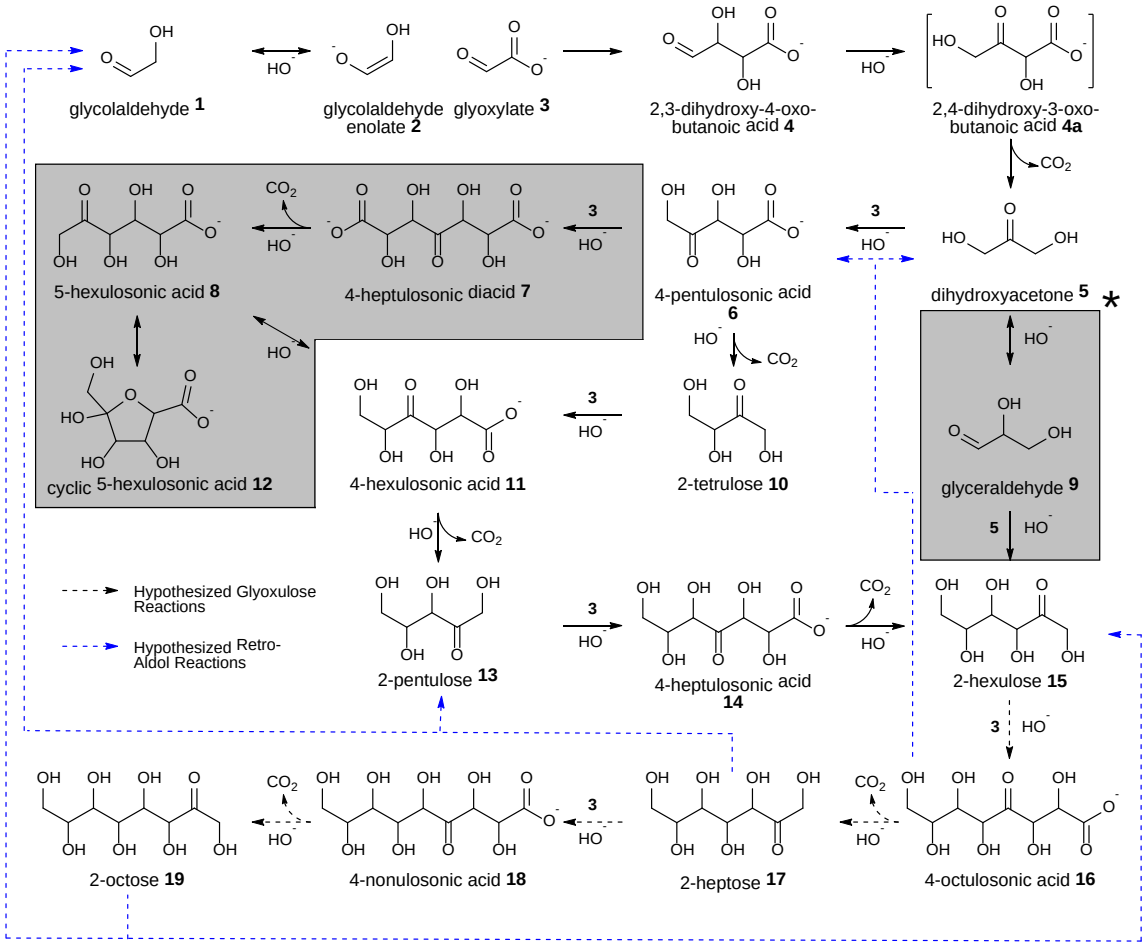


Figure 35 - Proposed overall reaction for the glyoxylose pathway of sugar formation. After initial seeding with a small amount of glycolaldehyde (or alternately any other ketose) glyoxylate is the only necessary addition to allow continued growth of linear sugars. The necessary steps for these transformations are illustrated in greatest depth for the converstion of glycolaldehyde to dihydroxyacetone, however the mechanism is fundamentally the same for every ketose. The first step is deprotonation of a terminal proton alpha to the carbonyl, this promotes tautomerization to the nucleophilic enolate (Conversion of 1 to 2, above). This enolate can then attack the aldehyde moiety of glyoxylate resulting in a γ-keto-ulosonic acid two carbons longer than the parent ketose (4). This γ-keto-ulosonic acid then undergoes carbonyl migration to the β-keto isomer(4a), which will readily decarboxylate yielding the next higher ketose (DHA, 5 in the case of glycolaldehyde). This series of reactions expected to recurse until species of chain length 8 are formed, at which point, while the forward glyoxylose reaction can still occur, competing retro aldol-reactions can be expected to feed back into the lower molecular weight reactants, ultimately resulting in a polydispersity centered around pentulose. Side reactions are shown above in gray. Double addition of glyoxylate to DHA is expected to yield heptulosonic acid and hexulosonic acids, which may be a sink for some of the reactive material. *Dimerization of DHA by tautomerization to glyceraldehyde and subsequent aldol reaction to hexose has been ruled out as a significant contributor due to DHA stability over weeks at reaction conditions.

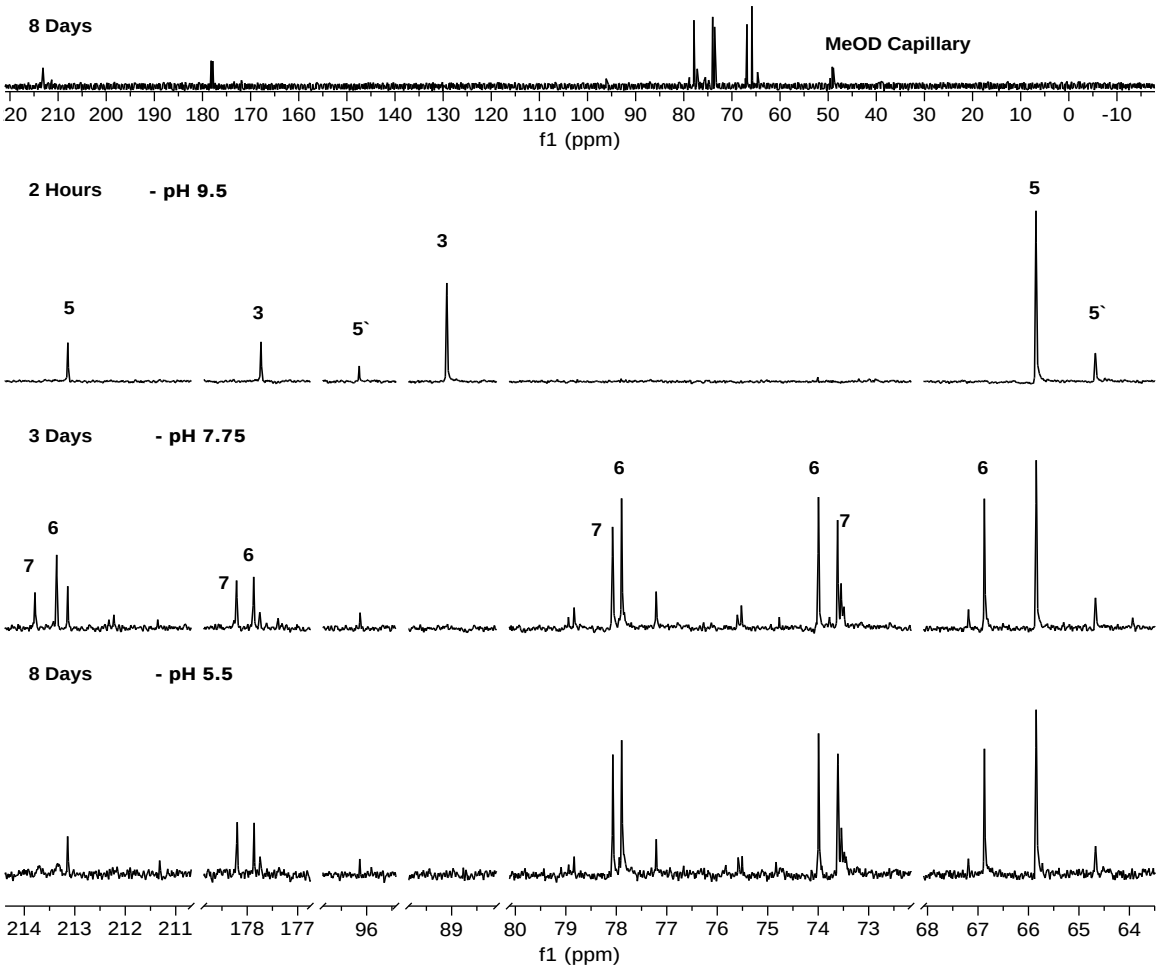


Figure 36 - Glyoxylose Transformation of Glyoxylate 3 and DHA 5 to Pent-4-ulosonic Acid 6 and Hept-4-ulosonic Di-Acid 7 The first investigated transformation in the glyoxylose reaction pathway is from DHA **5** to tetrulose **10**. Equal amounts of glyoxylate **3** and DHA **5** are reacted in aqueous NaOH with an initial pH of 9.5 at 30^{o}C for 8 days, and monitored by ^{13}C NMR spectroscopy. After two hours glyoxylate **3** (177.68 and 88.90 ppm), DHA **5** (213.07 and 65.65 ppm), and DHA monohydrate **5`** (95.96 and 64.46 ppm) are the main species observed. Five miniscule peaks were observed at 213.33, 177.87, 77.78, 73.87, and 66.76 ppm. These chemical shifts were consistent with carbons found in the functional groups present within pent-4-ulosonic acid **6**; ketone C=O, carboxylate C=O, two HO-CH, and HO-CH_2, respectively. After three days the signals associated with glyoxylate **3** are completely absent and peaks associated with DHA **5**, DHA monohydrate **5`**, and pent-4-ulosonic acid **6** were observed along with four new peaks. These four new peaks at 213.73, 178.19, 78.71, and 73.41 ppm were identified as hept-4-ulosonic diacid **7** which is formed when pent-4-ulosonic acid **6** reacts with glyoxylate **3**; we found that we could selectively form hept-4-ulosonic diacid **7** (two major diastereomers) by reacting DHA **5** with two equivalents of glyoxylate **3**. After eight days the peaks associated with pent-4-ulosonic acid **6** and hept-4-ulosonic diacid **7** remained unchanged, however an unknown process caused the pH to continue to drop. Additional experiments showed that addition of NaOH, increasing the pH, would prompt decarboxylation, causing the reaction to proceed to sugars.

While ^{13}C-NMR did not reveal proof of tetrulose **10** formation, it is a relatively low sensitivity analytical technique and may not reveal minor reaction products. To determine whether tetrulose **10** was present below the detection limit of our NMR analysis, direct inject electrospray ionized mass spectrometry was used to analyze the crude reaction mixture, revealing a mass peak consistent with tetrulose (Figure 37). Since the transformation of 4-oxo-pentulosonic acid **6** to 3-oxo-pentulosonic acid and subsequent decarboxylation could theoretically occur in the ionization process of the mass spectrometer, additional proof of tetrulose formation was sought by performing a nitrophenylhydrazine (NPH) derivatization[14, 15] of the reaction mixture and comparing to the derivative of an authentic tetrulose standard. These derivatized samples were compared using a liquid chromatography separation and ESI-MS. This comparison confirmed both by retention time and derivative mass to charge ratio, that the tetrulose was in fact produced in the reaction (Figure 38).

Compound	Molecular Ion	m/z	Compound	Molecular Ion	m/z
Glyoxylic Acid	[M-1]	73.2	Pentulosonic Acid	[M-1]	163.1
DHA	[M-1]	89.3	Hexulosonic Acid	[M-1]	193.0
Tetrulose	[M-1]	118.9	Heptulosonic Di-Acid	[M-2+Na]	258.9

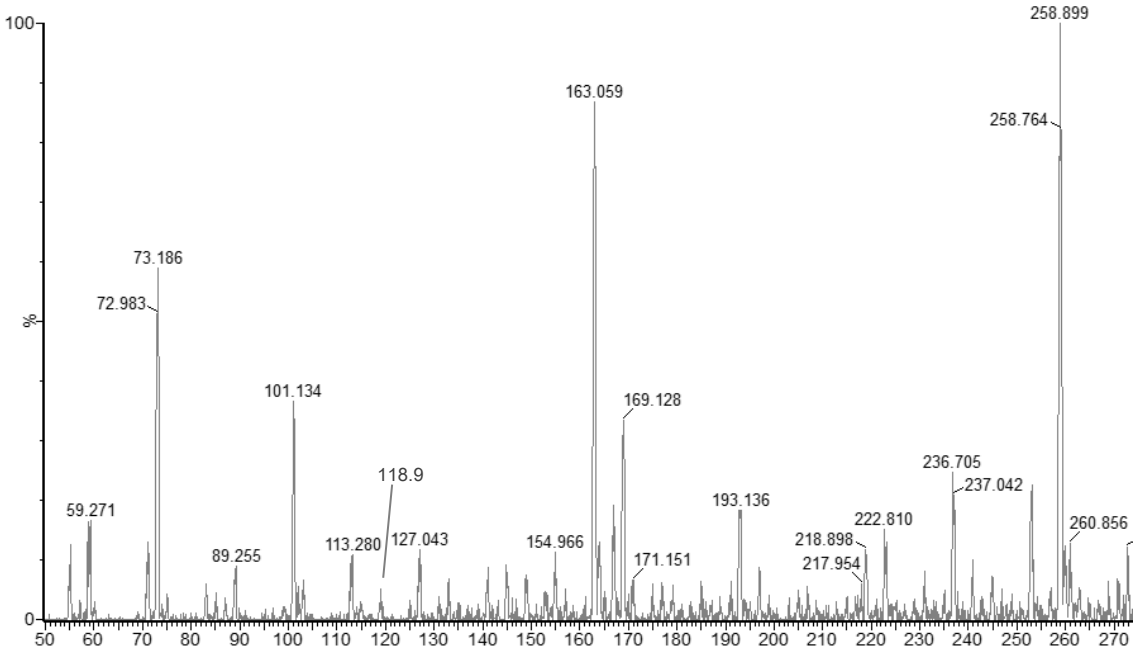


Figure 37 - MS Spectrum (ESI negative) of Glyoxylate 3 and DHA 5 Reaction and expected molecular peaks. Strong sigmals are observed for glyoxylic **3**, pentulosonic **6**, hexulosonic acids **11**, and heptulosonic diacid **7**, due to their more efficient ionization from carboxylate groups. Small peaks consistent with dihydroxyacetone **5** and tetrulose **10** are also observed.

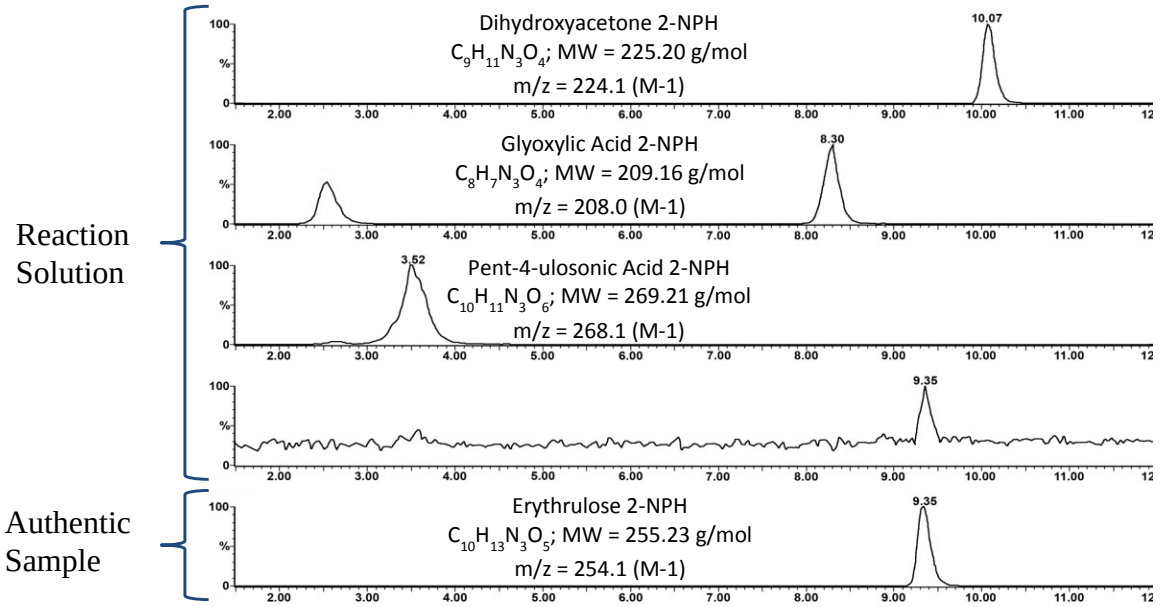


Figure 38 - Nitrophenylhydrazine Derivatization of DHA and Glyoxylate Reaction Mixture and Comparison to Tetrulose Standard The plot above are ion traces versus time obtained using LCMS. The signals correspond to the flux of ions of the desired mass. Retention times were compared against authentic standards. The NPH derivative of Erythrulose **10** is shown for illustrative purposes.

While our previous reaction showed that tetrulose was in fact produced from the reaction of glyoxylate **3** and DHA **5**, the continual readjustment of the pH may obscure important information about the reaction conditions necessary to effect this transformation. To remove some of the uncertainties associated with the above, we transitioned to a reaction of ^{13}C labeled reactants in a system buffered by phosphate at pH 8. The increased ease of NMR analysis of these ^{13}C labeled reactions allowed us to drop to reactant concentrations of 0.05M with our buffer at 0.5M total phosphate. This reaction was monitored by ^{13}C NMR and allowed to proceed for 7 days, at which time the pH had dropped to ca. 7.7. The NMR data revealed a greater extent of decarboxylation than the unlabeled (and higher concentration) reactions as well as peaks consistent with pentulosonic acid **6**, hexulosonic acid **8/11**, and heptulosonic diacid (Figure 39) **7**. Peaks

consistent with erythrulose **10** were tentatively identified, in addition to considerable (unreacted) DHA **5**. The abundance of C-C splitting made identification of further products difficult. Mass spectral comparisons were also made between the reaction of unlabeled DHA with labeled and unlabeled glyoxylate. These mass spectra clearly showed production of pentulosonic acid with the expected labelling, as well as tetrulose with a single enriched carbon, hexulosonic acid with three enriched carbons, and

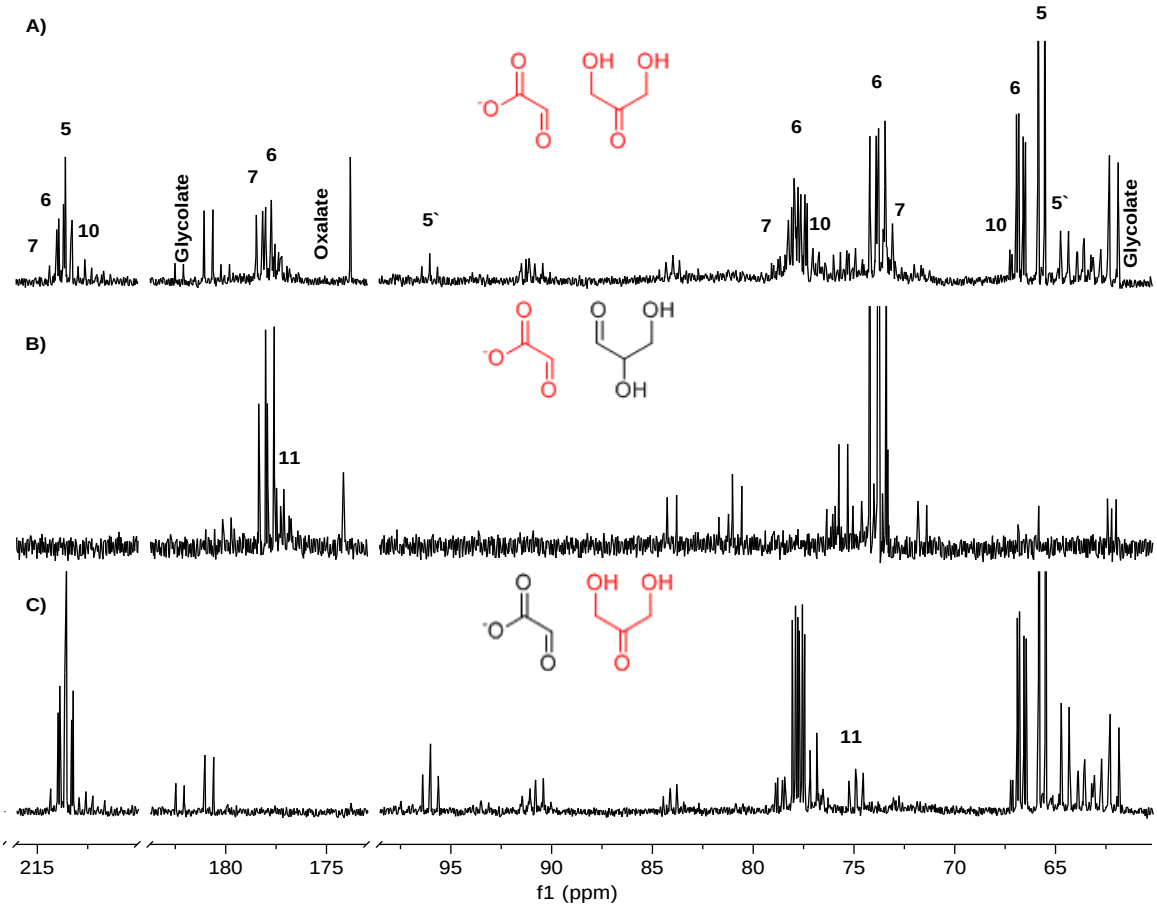


Figure 39 - ^{13}C NMRs of the Reactions of Glyoxylate and Dihydroxyacetone with Mixed ^{13}C Labeling After 1 Week. A) Reaction of 0.05M ^{13}C dihydroxyacetone and 0.05M ^{13}C glyoxylate in pH 8 0.5M phosphate. B) Reaction of 0.05M dihydroxyacetone and 0.05M ^{13}C glyoxylate in pH 8 0.5M phosphate. C) Reaction of 0.05M ^{13}C dihydroxyacetone and 0.05M glyoxylate in pH 8 0.5M phosphate. Labeling of the reactants is indicated in red. Not all peaks in the above NMR have been identified many are labeled impurities. Peaks consistent with considerable (unreacted) DHA **5** are present. Pentulosonic acid **6,** hexulosonic acid **8/11,** and heptulosonic diacid **7** are also firmly identified. Peaks associated with erythrulose **10** are also tentatively identified.

In an attempt to both somewhat simplify reaction spectra and investigate the effect of reaction stoichiometry, the reaction was repeated under in pH 8 phophate buffer with ^{13}C DHA **5**. In this case, however, unlabeled glyoxylate **3** was added in varying ratios at 1,2,5, and 10 equivalents with respect to DHA **5** and the reaction monitored over two weeks using ^{13}C NMR. Qualitative trends[1] of the depletion of dihydroxyacetone, and the production of pentulosonic acid **6**, hexulosonic acid **8/11**, heptulosonic diacid **7**, erythrulose **10**, and pentulose **13** are shown in Figure 40. These traces reveal that, as expected, DHA **5** is depleted more rapidly with increasing glyoxylate **3** concentration. Erythrulose **10** and pentulosonic acid **6** reach an initial maximum (which is less with increasing glyoxylate **3** concentration) and then are depleted as they convert to higher molecular weight species. Heptulosonic acid **7** and pentulose **13** increase with time in all cases except 10 equivalents glyoxylate **3**, in which case they reach a maximum and begin to decrease. Hexulosonic acid **8/11** is more difficult to analyze due to its potential to cyclize, multiple pathways of formation which obfuscate splitting paterns, and potential for multiple diastereomers, however, at least one linear form appears to be at a relatively constant low concentration with a small increase both with time and glyoxylate concentration. These observations indicate that as initially predicted availability of additional glyoxylate **3** is the sole factor in driving the reaction to higher molecular weights. The high accumulation of ribulose **13** in 10 equivalents glyoxylate **3** as compared to erythrulose is particularly interesting as it suggests the availability of a

[1] Trends are derived from integrals of an isolated peak of each species. NMRs are normalized by the integral of the methanol capillary. The largest integral per species is set to 1, with all other species analyzed by comparison.

cyclic form that prevents further reaction and, therefore, promotes accumulations, making this reaction more selective toward higher ketoses that can exist in cyclic forms.

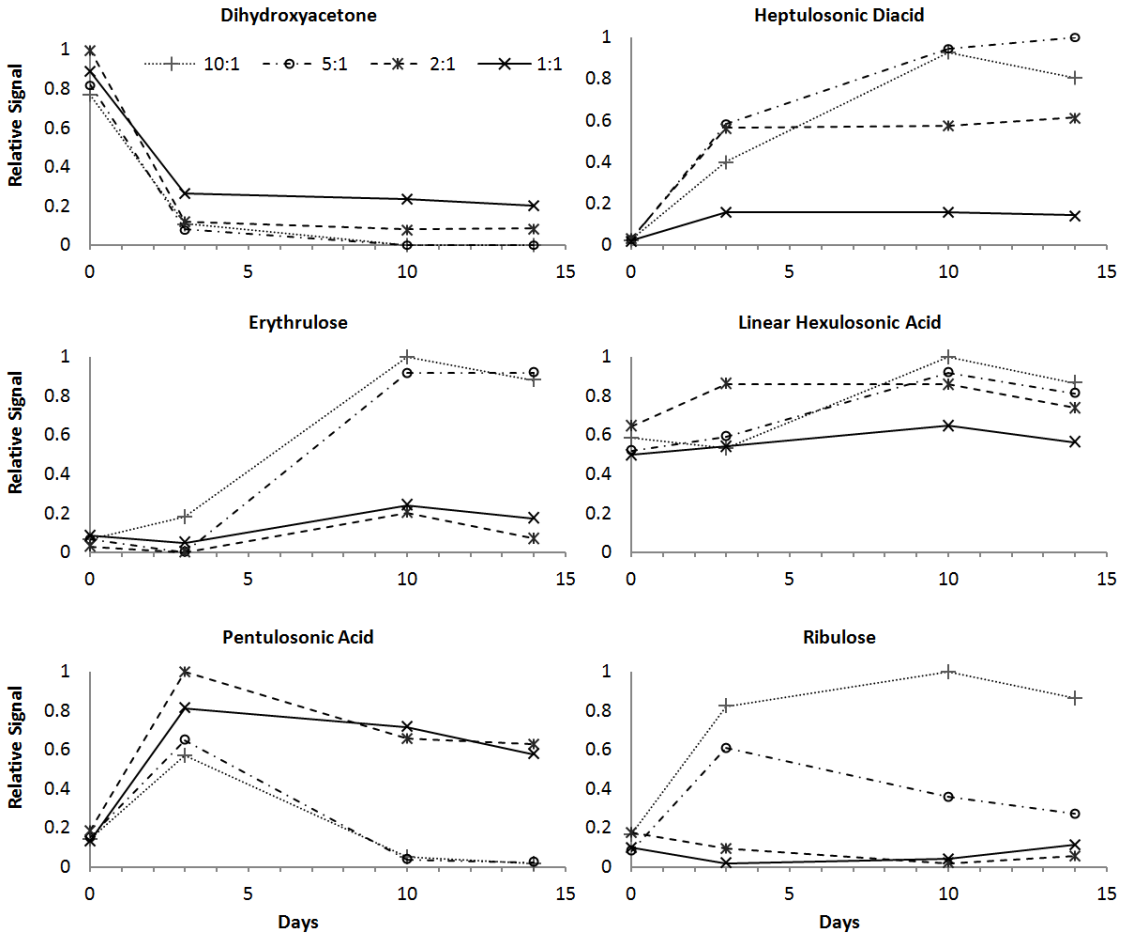


Figure 40 - Comparison of NMR Signals of Species of Interest in the Reaction of DHA and Glyoxylate. 0.05M ^{13}C Dihydroxyacetone was reacted with the specified equivalents of glyoxylate in 0.5M phosphate at pH 8. ^{13}C NMRs were collected after 3 hours, 4, 10, and 14 days. Comparison of traces shows that at higher equivalents of glyoxylate the reactants become bound in the ulosonic acid forms with decarboxylation products reacting rapidly with an additional equivalent of glyoxylate. Plots are for the investigation of trends only, and comparisons should not be made between species.

3.1.2 Reaction of Glyoxylate with Erythrulose

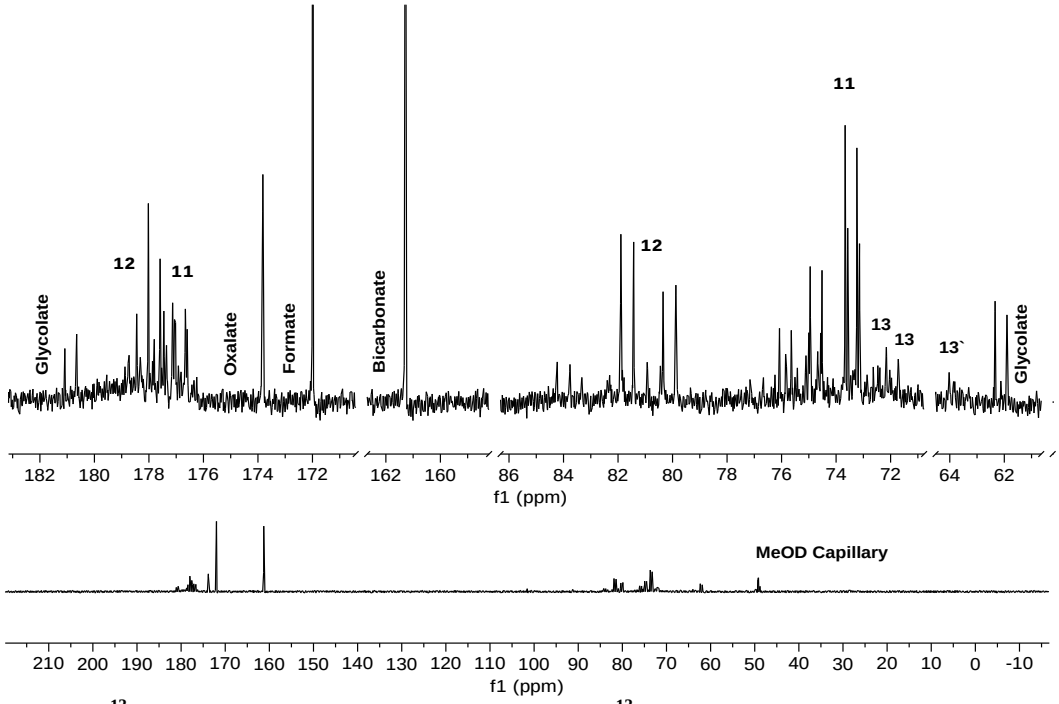


Figure 41 – ^{13}C NMR of the reaction of Erythrulose and ^{13}C Dilabeled Glyoxylate in pH 8 Phosphate Buffer after 1 Week. With this labeling scheme, only the peaks originating from glyoxylate are visible. Peaks identified with two diasteromers of hex-4-ulosonic acid **11** and two diasteromers of cyclic hex-5-ulosonic acid **12** are indicated. Additional peaks must correspond with linear hex-5-ulosonic acid **8** but assignments for this species are tenuous. A very large bicarbonate peak indicates that the glyoxylose reaction is occurring. Two peaks marked with **13** correspond to linear ribulose and xylulose. Small peaks marked with **13`** correspond to cyclic ribulose or xylulose.

To confirm that the ribulose noted in the previous experiment was in fact produced by the glyoxylose reaction of erythrulose **10**, a reaction of 0.05M erythrulose **10** and 0.05M ^{13}C glyoxylate **3** was conducted in 0.5M pH 8 phosphate buffer. After 1 week, investigation by ^{13}C NMR revealed peaks at 176.5 and 73.4 ppm consistent with two diastereomers of hexulosonic acid **11**, containing two labeled carbons originating from glyoxylate **3**, being formed. A large bicarbonate peak at 161 ppm indicates that decarboxylation of this hexulosonic acid **11** is occurring. This is further substantiated by singlet peaks at 71.9 and 72.1 ppm that are consistent with the terminal carbons of ribulose and xylulose **13**. A singlet at 64 ppm is also consistent with the terminal carbon

a cyclic pentose **13`**. Peaks were also observed that were consistent with the cyclic form **12** of 5-hexulosonic acid **8**. This cyclic hexulosonic acid **12** is consistent with the addition of glyoxylate to erythrulose followed by carbonyl migration in the "wrong" direction to the five position, rather than the three position which would result in decarboxylation. This ring form can theoretically open and undergo carbonyl migration again, ultimately resulting in the 3-hexulosonic acid which will decarboxylate, however we have no evidence for or against this occurring in practice.

The reaction of 0.1M erythrulose and ^{13}C labeled and unlabeled glyoxylate was also conducted at these conditions and monitored by LCMS after one week. These reactions clearly showed dilabeled hexulosonic acid formation as well as a small amount of trilabeled heptulosonic acid, which indicates the formation of monolabeled pentulose, consistent with the proposed mechanism. Pentulose was not observed directly, however, the employed LCMS method performs poorly in the ionization of sugars.

Formation of the formate byproduct observed in these reactions may be attributable to dehydration and retro-aldol reactions often observed in Maillard reactions[16]. This mechanism can also be invoked to explain the acidification of the reaction mixture which is not explained by the DHBA producing reaction.

3.1.3 Reaction of Glyoxylate with Ribulose

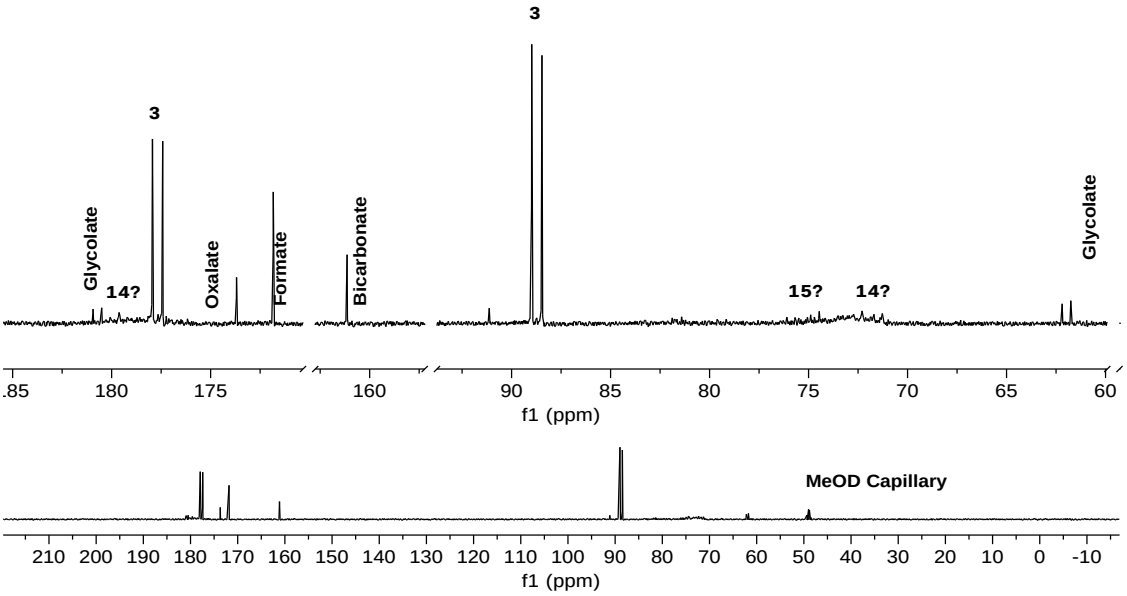


Figure 42 – ^{13}C NMR of the reaction of Ribulose and ^{13}C Dilabeled Glyoxylate in pH 8 Phosphate Buffer After 1 Week. With this labeling scheme, only the peaks originating from glyoxylate are visible. The most prominent peaks in Additional peaks must correspond with linear hex-5-ulosonic acid **8** but assignments for this species are tenuous. A very large bicarbonate peak indicates that the glyoxylose reaction is occurring. Two peaks marked with **13** correspond to linear ribulose and xylulose. Small peaks marked with **13`** correspond to cyclic ribulose or xylulose.

The accumulation of ribulose in the reaction of DHA and ten equivalents of glyoxylate suggests that a reaction begun with ribulose should progress to a significantly smaller degree than that of lower ketoses. This was supported by the results of a reaction of ribulose and 13C glyoxylate under equivalent reaction conditions. In this reaction, 0.05M ribulose was combined with 0.05M ^{13}C glyoxylate in 0.5M pH 8 phosphate buffer. After 1 week, the major signals were those of unreacted glyoxylate indicating that ribulose is significantly less reactive than DHA and erythrulose, and lending strength to the hypothesis that this lower reactivity would cause five carbon species to preferentially accumulate in the glyoxylose pathway. This lower reactivity is not zero reactivity. Peaks around 179 ppm and 74 ppm are consistent with expected peaks for the 7 carbon ulosonic

acid **14** containing two labeled carbons originating from glyoxylate **3**. Peaks in the mid 70 ppm range are consistent with the expected shifts of the single labeled carbon in a hexose species originating from this pathway. The presence of quite a few small peaks in these regions can be attributed both to the likely production of multiple diastereomers, and the potential for an equilibrium relationship between the linear, furanosyl, and pyranosyl forms of these species.

The reaction of 0.1M ribulose and ^{13}C labeled and unlabeled glyoxylate was also conducted at these conditions and monitored by LCMS after one week. These reactions clearly showed dilabeled heptulosonic acid formation. No other products were observed, consistent with the slow pace of this reaction. The reaction of 0.1M fructose and ^{13}C labeled and unlabeled glyoxylate was also investigated, revealed formation of dilabeled octulosonic acid.

The continued reactivity of the glyoxylose pathway beyond five carbon sugars is important to the viability of the reaction pathway. While it is of course desirable to accumulate five carbon sugars, due to their importance in modern genetic material, some continued reactivity allows for the possibility of autocatalysis. Longer sugars are more susceptible to retro-aldol reactions, a fact exploited by nature for the synthesis of four carbon sugars. In the pentose phosphate pathway, sedoheptulose 7-phosphate is fragmented to form erythrulose 4-phosphate, and glyceraldehyde 3-phosphate[17]. This same fragmentation to lower sugars can occur in non-phosphorylated sugars under hydrothermal conditions[18]. Each of these smaller species can tautomerize to a ketose form and react with glyoxylate allowing further production and accumulation of pentuloses.

3.1.4 Reaction of Glyoxylate with Glycolaldehyde

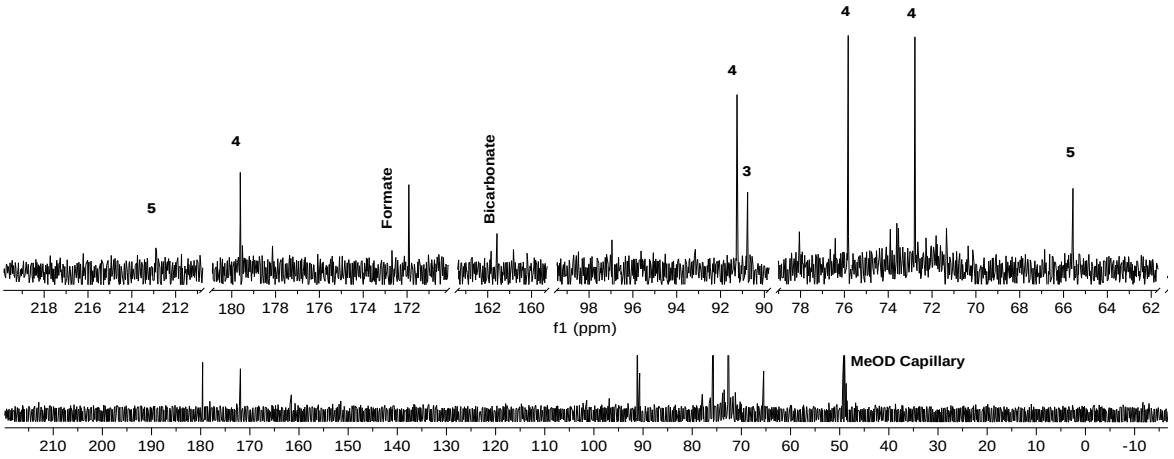


Figure 43 – ^{13}C NMR of the Reaction of 0.5M Glycolaldehyde and 0.5M Glyoxylate adjusted to pH 8.5 daily for 6 days. The reaction of glyoxylate and glycolaldehyde yields dihydroxybutanoic acid **4** as the major product after 6 days. A small amount of bicarbonate and signals corresponding to dihydroxyacetone indicate that carbonyl migration and decarboxylation did occur. Small peaks in the 70-80ppm range also suggest that some DHA reacted further with glyoxylate creating additional glyoxylose products.

Thus far we have addressed the reaction of ketoses with glyoxylate; however it should also be possible to begin the glyoxylose reaction with the reaction of glycolaldehyde **1** and glyoxylate **3**. Glycolaldehyde **1** is a more attractive starting species due to its greater simplicity, and assumed prebiotic abundance [19-21]**.** Due to the potential for further interactions with glycolaldehyde we chose to investigate this reaction under a wider variety of conditions similar to the previously discussed reaction of DHA **5** and glyoxylate **3**.

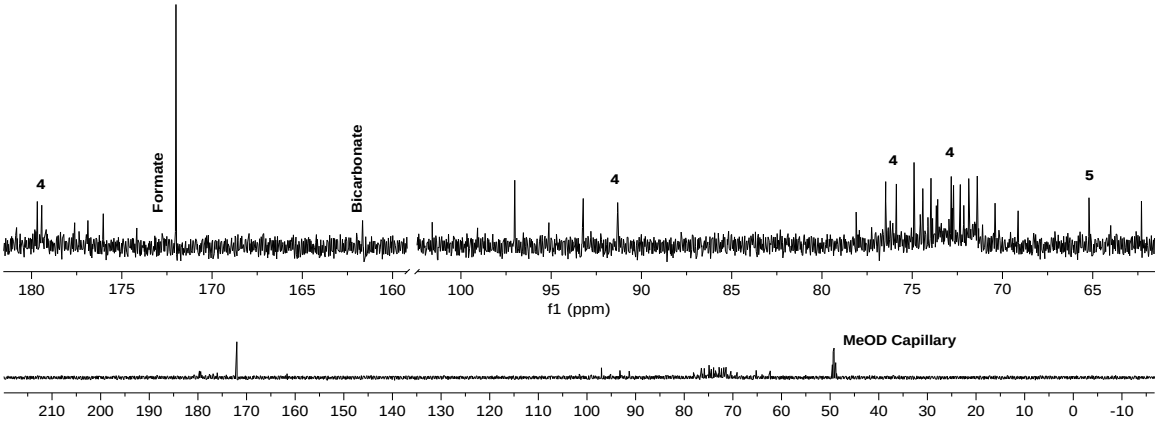


Figure 44 – ^{13}C NMR of the Reaction of 0.5M Glycolaldehyde and 0.5M Glyoxylate adjusted to pH 8.5 daily for 12 days. After 12 days, the reaction of glyoxylate and glycolaldehyde no longer contains any trace of the initial glyoxylate **3**. Peaks consistent with dihydroxybutanoic acid **4** persist, however to these peaks are added 20+ additional peaks indicating that the glyoxylose reaction has proceeded further. A small amount dihydroxyacetone **5** remains. The continual acidification causes conversion of the bicarbonate to CO_2 suppressing that signal.

Initial investigations of the reaction of glycolaldehyde **1** and glyoxylate **3** were conducted by combining 0.5 M of each reactant in LCMS grade water and adjusting the initial pH to near 8.5 by addition of NaOH. After 1 day, the pH of the reaction had decreased to ca. 4.5 (a much larger drop than observed in the reaction of DHA and glyoxylate). By ^{13}C NMR some conversion was observed, however glyoxylate and glycolaldehyde persisted as major signals. The reaction pH was readjusted to 8.5 daily by addition of NaOH and monitored by ^{13}C NMR. After 6 days (Figure 43), the reaction was found to form primarily the semialdehyde of tartaric acid(2,3-dihydroxy-4-oxo-butanoic acid; DHBA **4**). Peaks at 212 ppm (DHA carbonyl) and 161 ppm (bicarbonate) were also observed indicating decarboxylation of DHBA to dihydroxyacetone **5,** potentially explaining the pH change as carbonic acid is produced. After 12 days (Figure 44) the ^{13}C NMR of the reaction becomes quite complex exhibiting peaks likely belonging to pentulosonic acid **6** and heptulosonic diacid **7** and small peaks in the 90-105ppm range which could indicate anomeric carbons of a pentulose **13** or hexulose **15**

product. Peaks in this range are instructive as the spectral region from 68-78ppm which corresponds to -HCOH- signals contains no fewer than 25 peaks making a comprehensive analysis of this region difficult.

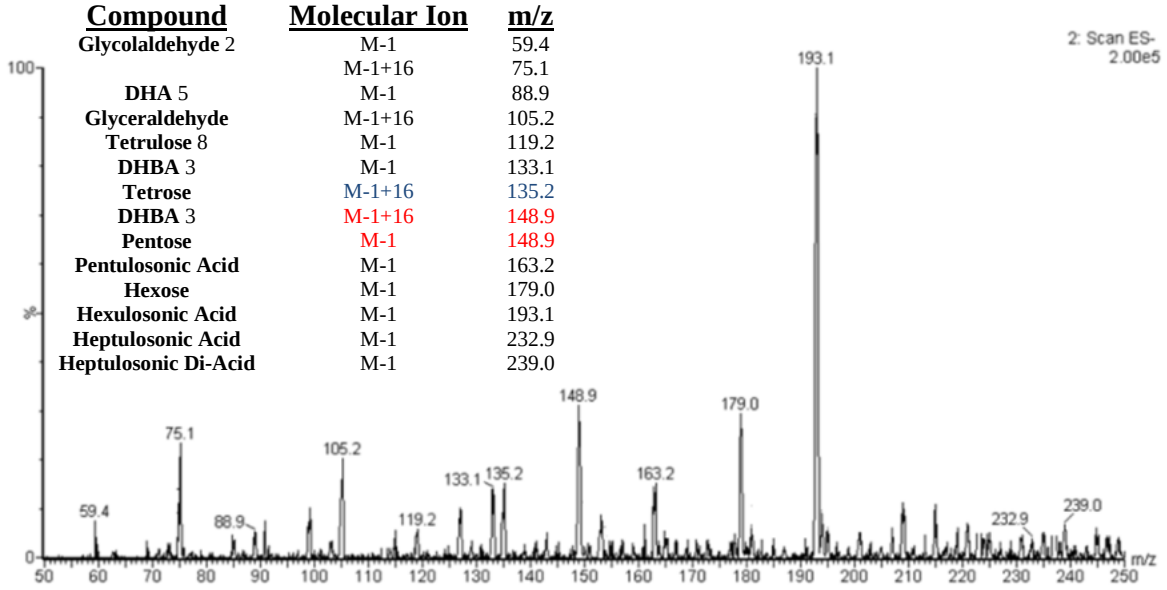


Figure 45 – MS Spectrum (ESI negative) of Glyoxylate 1 and Glycolaldehyde 2 Reaction The ESI negative MS spectrum yielded m/z that were consistent with glycolaldehyde 2 (59.4, M-1), DHBA 3 (133.1, M-1), and DHA 5 (88.9, M-1) indicating to us that the three steps of the glyoxylose reaction had occurred; Aldol addition, C=O migration, and decarboxylation. Also present in this spectrum are m/z that are consistent with pentulosonic acid (163.2, M-1) and tetrulose 8 (119.2, M-1) which are formed in the sequence of the glyoxylose reaction; DHA 5 reacts with glycolaldehyde 2 (59.4, M-1), to form pent-4-ulosonic acid 6 which undergoes C=O migration and loss of CO_2 to form tetrulose 8. Under basic conditions in ESI negative two m/z ratios were observed for the triose species; the ketose (DHA 5) ionized as expected giving the M-1 molecular ion at 88.9 while the aldose (glyceraldehyde) underwent oxidation from an aldehyde to a carboxylate producing a molecular ion at 105.2, the M-1+16 anion; this oxidation was also observed for glycolaldehyde 2 (59.4, M-1, 75.1, M-1+16) and tetrose (119.2, M-1; 135.2, M-1+16. The presence of both the aldose [5] and ketose forms of the trioses provides the opportunity for these two to react with each other and form hexose (179, M-1) in a side reaction [6]. The other m/z identified in the MS spectrum are the result of the reaction of pent-4-ulosonic acid 6 with glyoxylate 1 to form hept-4-ulosonic di-acid 13 (239.0 M-1) which undergoes C=O migration and decarboxylation to form hex-5-ulosonic acid 15 (193.1, M-1).

A mass spectral analysis was also conducted of this reaction mixture, yielding data consistent with pentuloses and hexuloses having been produced. The crude reaction mixture was diluted in LCMS water and analyzed by infusion LCMS with ESI- ionization. Peaks were observed at 119.2, 148.9 and 179.0 m/z corresponding to the [M-

1] peaks of four, five and six carbon sugars. Mass to charge ratios were also observed corresponding to pentulosonic, hexulosonic, and heptulosonic acids; the sugar acid intermediates that precedes these erythrulose, pentulose, and hexulose species, respectively, in the glyoxylose pathway.

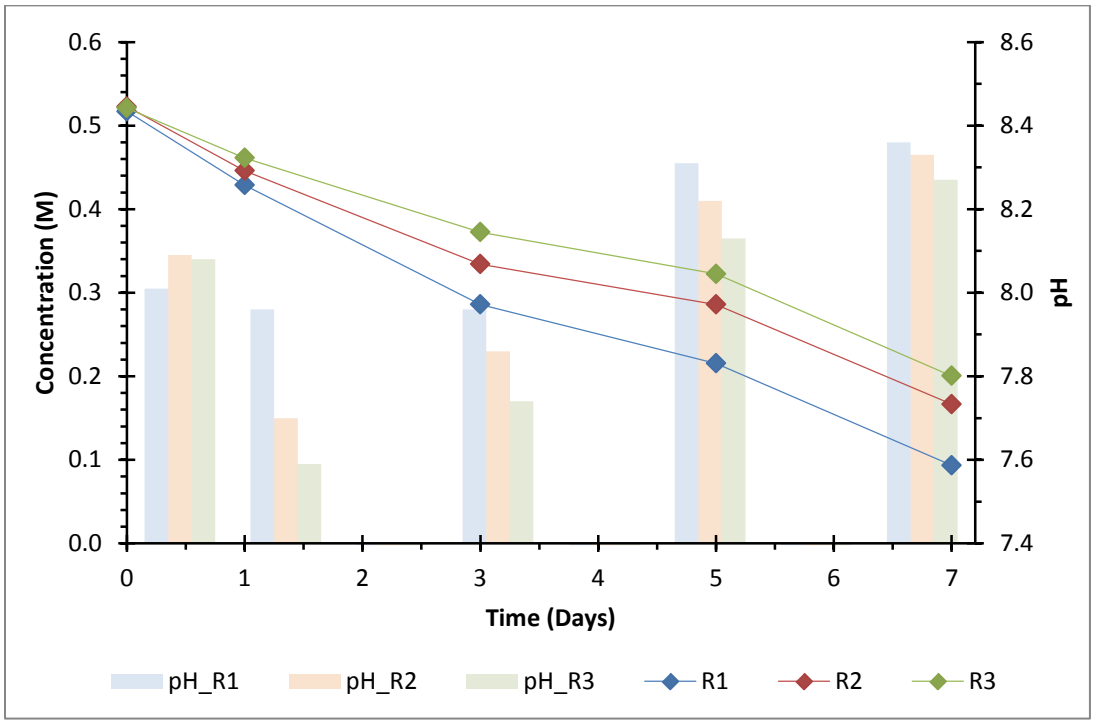


Figure 46 – Concentration of Glyoxylate with Time and pH in Three Reactions of 0.5M Glyoxylate and 0.5M Glycolaldehyde. Concentration of glyoxylate as determined by nitrophenylhydrazine derivatization, and analysis by LCMS. pH of each of the reactions after being adjusted for the given day are shown. Adjustments were only conducted on the indicated days. There appears to be a correlation between pH at the beginning of each period and the amount of glyoxylate consumed in the following days.

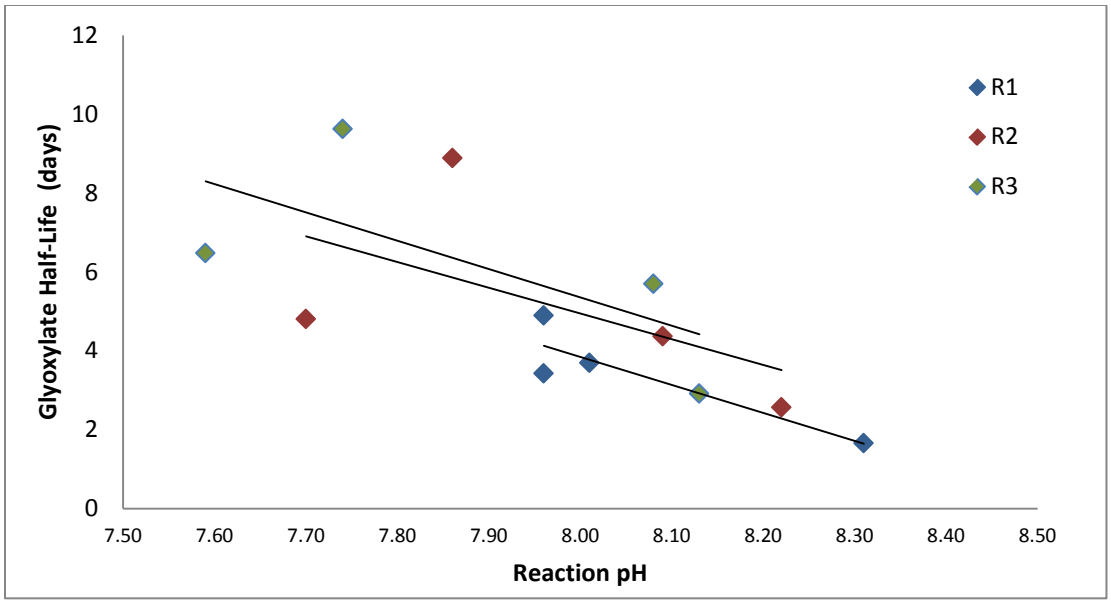


Figure 47 – Discrete Reaction Half-Life of Glyoxylate in Three Reactions of 0.5M Glyoxylate and 0.5M Glycolaldehyde as a function of pH. Half-lives calculated from the data presented in Figure 46 using the equation:

$$t_{1/2} = \frac{(t_2 - t_1) \cdot ln(\frac{1}{2})}{ln(\frac{C_2}{C_1})}$$

Where:

$t_{1/2}$ = the reaction half-life
t_1 = the time of the given concentration measurement
t_2 = the time of the subsequent concentration measurement
C_1 = the glyoxylate concentration at t_1
C_2 = the glyoxylate concentration at t_2

The above figure supports the relationship between pH and reaction rate. This data indicates that increasing the pH from 7.7 to 8.3 results in approximately quadruple the reaction rate.

Replicates were conducted of the reaction of 0.5M glycolaldehyde and 0.5M glyoxylate; however, differences in pH were introduced during the course of adjustment. Interestingly, significant differences in the amount of glyoxylate consumed were noted based on differences in pH of just 0.2. Figure 46 shows the disappearance of glyoxylic acid as the reaction progresses as well as the pH to which the reaction was adjusted each time an aliquot was removed for analysis. This is consistent with the equilibrium between glycolaldehyde **1 a**nd the glycolaldehyde enolate **2** being pH dependent[22] [16] and the reaction rate being driven by enolate concentration. Figure 47 further illustrates

this trend by showing the effect of pH on the half-life of glyoxylate in these reactions, demonstrating a linear decrease in half-life (thus faster reaction rate) as pH increases. This trend is consistent with reaction kinetics which can be approximated as first order in both glyoxylate and hydroxide concentration; this entails assuming that glyoxylate reacting with glycolaldehyde is the dominant reaction ($C_{glycolaldehyde} >> C_{higher\text{-}ketoses}$), or that the reaction rate of glyoxylate with glycolaldehyde is the rate controlling step during this portion of the reaction.

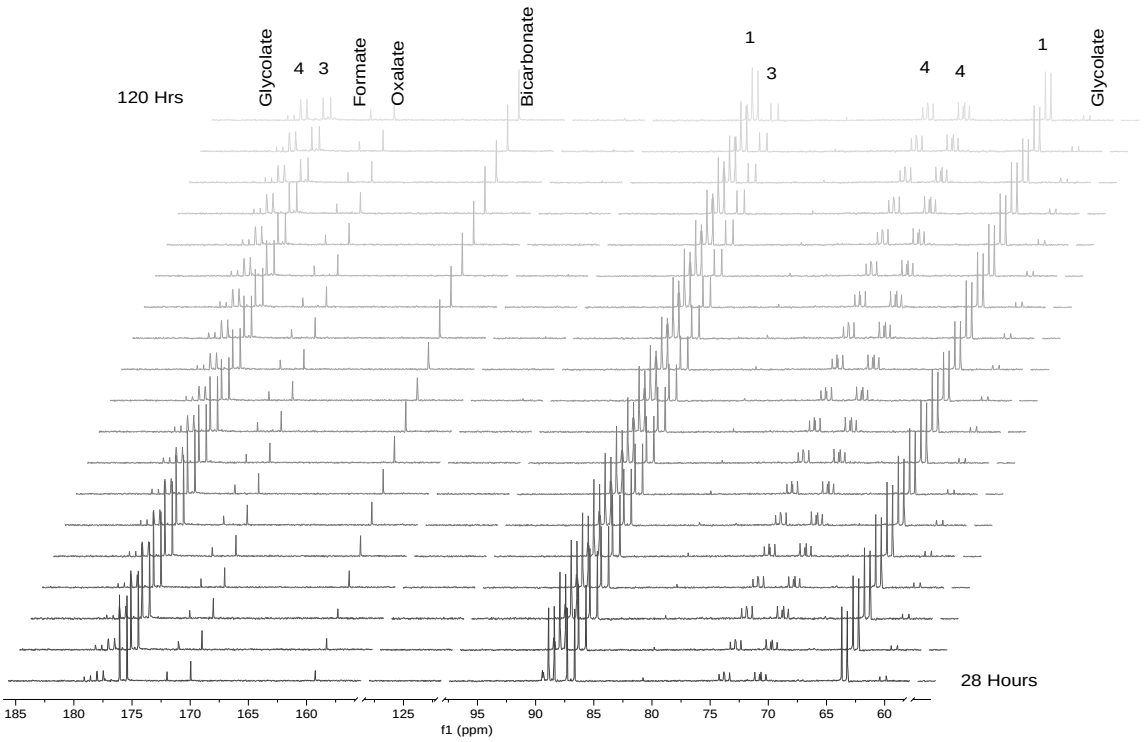


Figure 48 – Quantitative ^{13}C NMR Time Course of the Reaction of 0.05M ^{13}C Glycolaldehyde and 0.05M ^{13}C Glyoxylate in 0.5M pH 8 Phophate buffer from 1 to 5 Days. These NMRs clearly show production of 2,3-dihydroxy-4-oxo-butanoate **4** and bicarbonate, while glycolaldehyde **1** and glyoxylate **3** deplete. Production of DHA and higher products is evidenced by the enhanced depletion of glyoxylate relative to glycolaldehyde and production of bicarbonate; however, the only direct evidence of these products in the NMR is some raising of the baseline in the 65-80 and 175-180 ppm ranges. Due to low individual concentrations from diastereomerism and reduction of signal intensity from carbon-carbon splitting from ^{13}C labeling, this is not entirely unexpected.

To obtain more information about the kinetics of this reaction, 0.05M ^{13}C glyoxylate and 0.05M ^{13}C glycolaldehyde were reacted in a phosphate buffer with concentration of 0.5M and pH of 8. After a one day incubation period, the reaction was monitored over time using a quantitative ^{13}C time course from 1 to 4 days reaction time (Figure 48). This technique showed apparent first order kinetics in glyoxylate and glycolate consumption based on integrated peaks for each species. Signals consistent with conversion to DHBA were observed along with a marked increase in bicarbonate which must correspond to a conversion to dihydroxyacetone. Dihydroxyacetone was not directly observed at these concentrations, however this is not surprising as it is more reactive with residual glyoxylate than glycolaldehyde and is likely to undergo further reaction to pentulosonic diacid, which can react with another glyoxylate or decarboxylate to erythrulose. The diastereomers of these products, the carbon-carbon splitting, and the low number of scans in this technique all combine to effectively mask NMR signals from higher products.

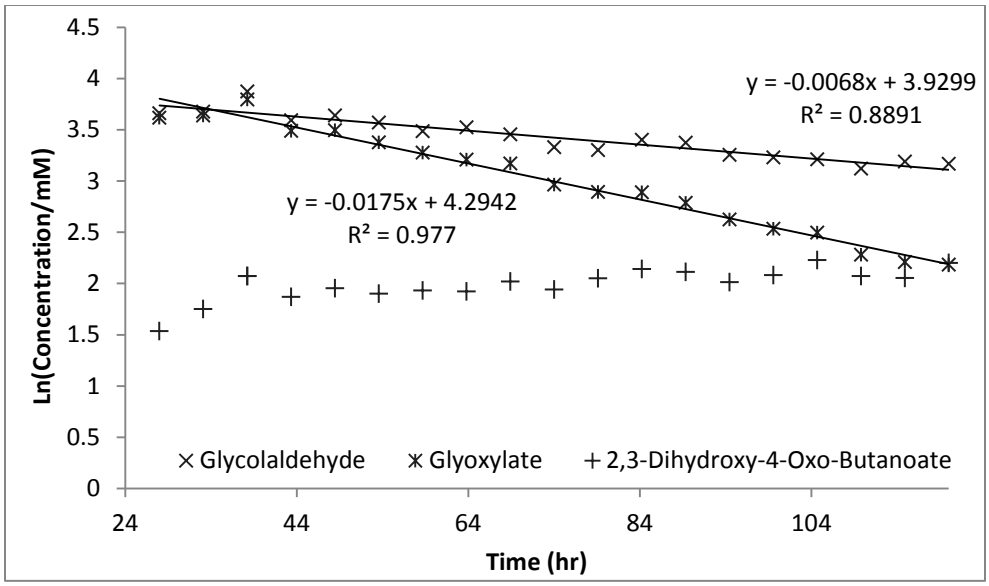


Figure 49 – Rate Profiles of Glycolaldehyde, Glyoxylate, and 2,3-Dihydroxy-4-Oxo-Butanoate. Obtained from the reaction depicted in Figure 48. The kinetic data shown here depicts pseudo first order kinetics for both glycolaldehyde **1** and glyoxylate **3**. DHBA **4** is revealed to be at a near steady state concentration. The rate of glyoxylate **3** consumption is slightly more than double that of glycolaldehyde **1**. This is consistent with the steady state concentration of DHBA **4**. For such a steady state to exist, for every molecule of DHBA **4** produced, one must convert to dihydroxyacetone **5**. DHA is quite reactive and can thus react immediately with an additional equivalent of glyoxylate to form pentulosonic acid **6**, thus explaining twice the glyoxylate consumption as related to glycolaldehyde. The additional reactivity can be explained by potential conversion of pentulosonic acid **6** to heptulosonic acid **7** and consumption of glyoxylate by the Cannizzaro reaction.

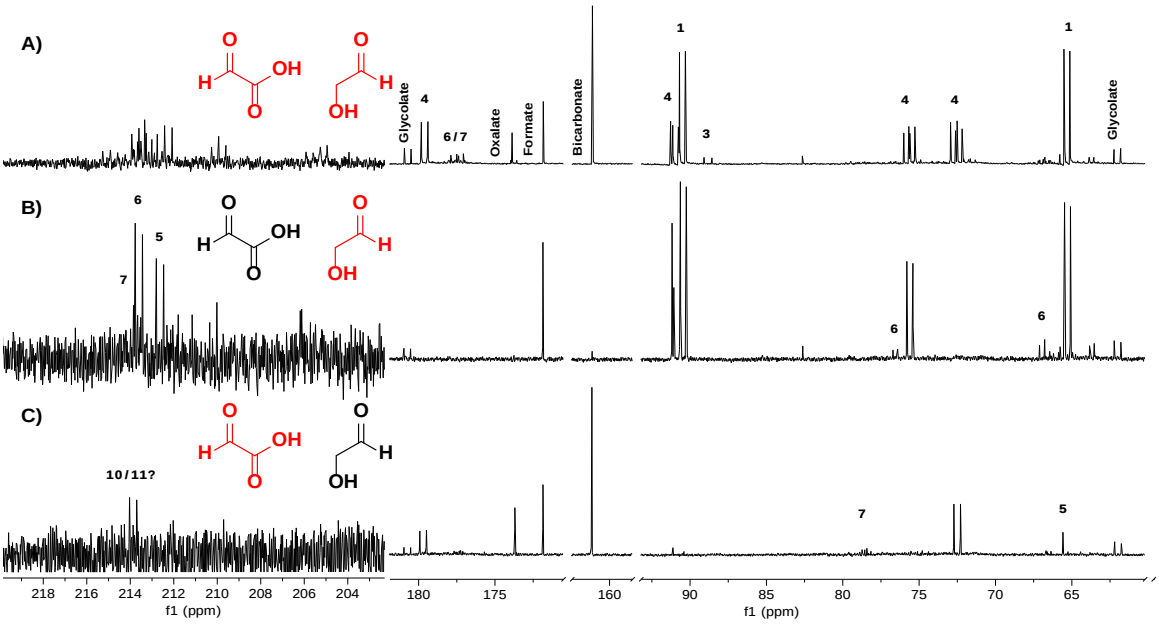


Figure 50 – ^{13}C NMRs of the Reactions of Glyoxylate and Glycolaldehyde with Mixed ^{13}C Labeling After 1 Week. A) Reaction of 0.05M ^{13}C glycolaldehyde **1** and 0.05M ^{13}C glyoxylate **3** in pH 8 0.5M phosphate. B) Reaction of 0.05M ^{13}C glycolaldehyde **1** and 0.05M glyoxylate **3** in pH 8 0.5M phosphate. C) Reaction of 0.05M glycolaldehyde **1** and 0.05M ^{13}C glyoxylate **3** in pH 8 0.5M phosphate. Labeling of the reactants is indicated in red. Not all peaks in the above NMR have been identified many are labeled impurities. Peaks consistent with considerable (unreacted) glycolaldehyde **1** are present. A small amount of unreacted glyoxylate **3** also remains. 2,3 dihydroxy-4-oxo-butanoate **4** is firmly identified. Peaks associated with pentulosonic acid **6** and heptulosonic diacid **7** can be seen clearly in the mixed labeled reactions. A peaks associated with erythrulose **10** or hexulosonic acid **11** is tentatively identified in NMR C) based on these being the first two species which could have a labeled ketone originating from glyoxylic acid.

To further simplify the analysis of this reaction, replicates were conducted using mixed labeled reactants. Using only one reactant with ^{13}C labeling effectively eliminates all signals from the unlabeled, reducing them to <4% of the total signal and also greatly simplifies the carbon-carbon splitting. These reactions were conducted using 0.05 M in each reactant, a 0.5M phosphate buffer at initial pH of 8 using ^{13}C labeling as indicated in Figure 50. The reactions were allowed to proceed for 1 week at which time the pH was ~7.6. When these reactions are compared to the fully labeled reaction, peak identities are immediately confirmed. A singlet at 66ppm in the labeled glyoxylate reaction must correspond to the production of dihydroxyacetone, just as a doublet at 213 ppm indicates the reaction has progressed through erythrulose (or potentially even to hexulose), the

first species in which a carbonyl originating from glyoxylate could occur (glyoxylate, itself, under these highly basic conditions exihibits only a hydrate signal). Further differences in the apparent peaks help to confirm the proposed reaction mechanism. Among these differences are: presence of carbonate (162 ppm) and dissolved CO_2 (126 ppm) peaks solely in the samples which contain labeled glyoxylate confirming decarboxylation taking place through the proposed mechanism; conversion of the DHA peak at 66ppm from a doublet to singlet when going from labeled glycolaldehyde to labeled glyoxylate; conversion from triplet to doublet to absent of the DHA carbonyl peak at 212 ppm when going from all labeled, to labeled glycolaldehyde, to labeled glyoxylate. While further monitoring of these reactions does reveal a gradual decrease in the amount of DHBA present in solution, no higher products were identified. This is likely due to a lack of free glyoxylate to promote chain growth.

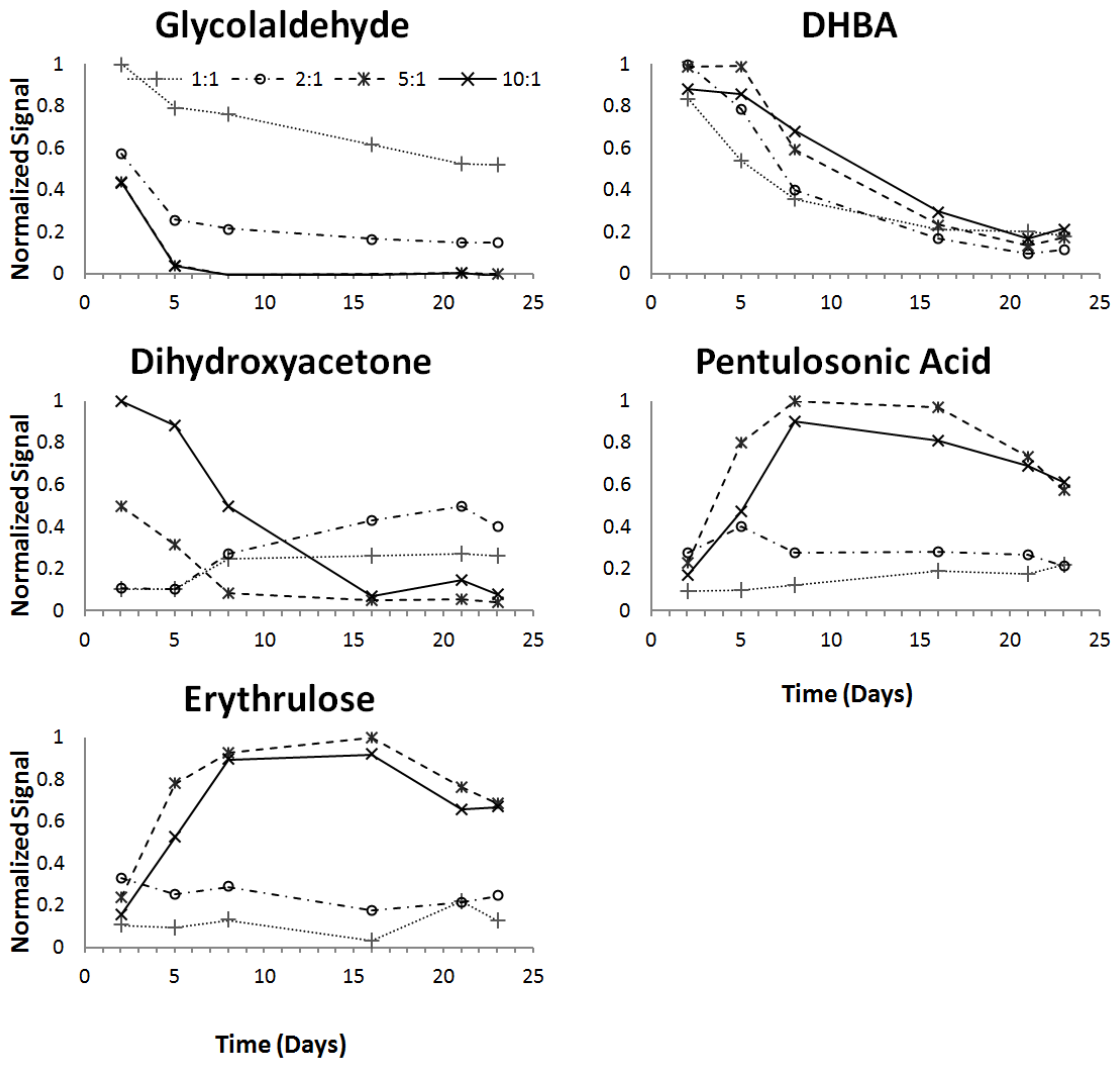


Figure 51 – Comparison of NMR Signals of Species of Interest in the Reaction of DHA and Glyoxylate. 0.05M ^{13}C Glycolaldehyde was reacted with the specified equivalents of glyoxylate in 0.5M phosphate at pH 8. ^{13}C NMRs were collected after 2, 5, 8, 16, 21, and 23 days. Comparison of traces shows that the rate of glycolaldehyde depletion increases with the number of equivalents of glyoxylate. DHBA hits an early maximum and decarboxylates at a relatively constant rate regardless of glyoxylate concentration. Maxima of both pentulosonic acid and erythrulose increase with the number of equivalents of glyoxylate. Plots are for the investigation of trends only, and comparisons should not be made between species.

To attempt to force the reaction to even longer chain lengths, we conducted a study of the effect of excess glyoxylate on the reaction. Reactions were conducted using 0.05M 13C glycolaldehyde in a 0.5M phosphate buffer at pH and from 1 to 10 equivalents of unlabeled glyoxylate. The reactions were monitored using 13C NMR for

three weeks. As would be expected, more equivalents of glyoxylate lead to greater consumption of glycolaldehyde. After three weeks the reaction with a single equivalent of glyoxylate showed a large concentration of glycolaldehyde remaining, and only a small peak at 96 corresponding to the hydrate of dihydroxyacetone. By contrast, the NMR with ten equivalents of glyoxylate shows a large peak at 96 ppm after 2 days which only decreases from there, indicating formation of pentulosonic acid. This ten to one reaction also shows peaks consistent with the formation of heptulosonic diacid, as well as many -HCOH-, and -H2COH peaks in the 60s and 70s consistent with 5 and 6 carbon sugar formation. There is a marked increase in complexity from the 1:1 to 10:1 reaction, consistent with the proposed reaction mechanism, only requiring a small amount of glycolaldehyde to reach relatively high chain length sugars.

A separate reaction of 0.1M glycolaldehyde with 0.1M glyoxylate in a pH 8 phosphate buffer was also conducted and monitored by LCMS with quantitation conducted using an MRM method tuned to ribulose. This method revealed pentulose yields as early as the second day of reaction time. Results can be seen in Figure 52 and are in generally good agreement with kinetic and compositional trends as observed by NMR.

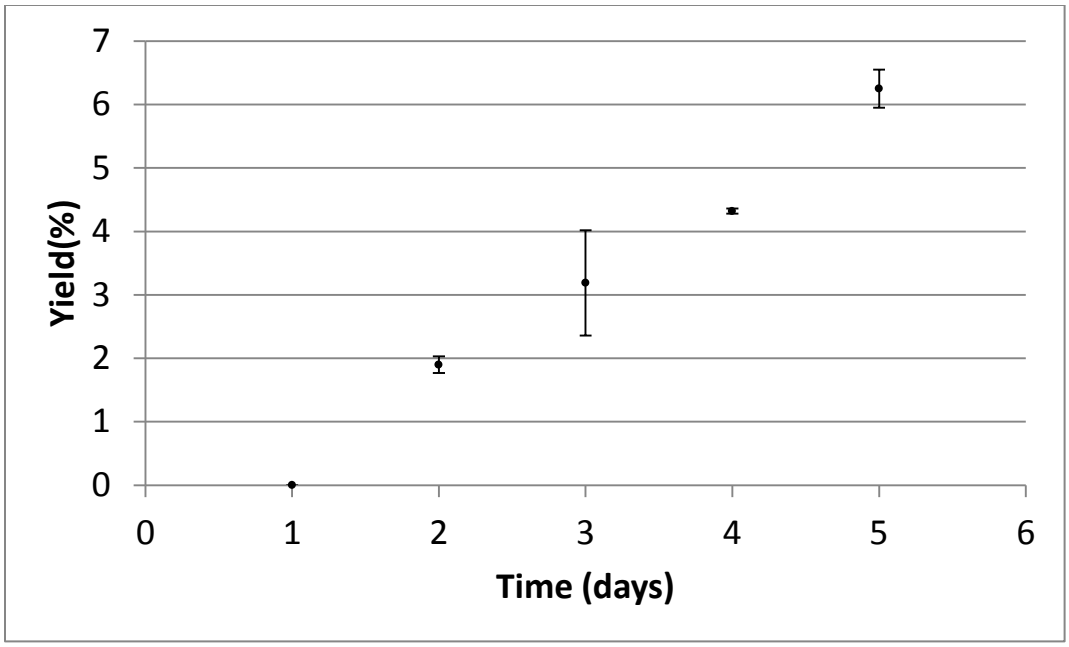


Figure 52 – Yield of Ribulose from Glyoxylose Reaction Over Time. Yields of ribulose over time from reaction of 0.1M glycolaldehyde with 0.1M glyoxylate in a pH 8 phosphate buffer. Yield calculated with respect to glycolaldehyde.

3.1.5 Summary

In none of the reactions discussed above do we observe evidence of branched sugar formation (e.g. proton-less carbons being split by three adjacent species in NMR). This can be attributed to the carbonyl migration step in the glyoxylose pathway. If a branched species is formed, carbonyl migration to the beta position to the acid cannot occur, blocking further reactivity. Ultimately retro-aldol reaction to free glyoxylate should occur, causing the linear species to outcompete the branched (Figure 53). This is aided by the statistical increase in likelihood of removing one of the two terminal protons as compared to removing the single proton at the 3 position of a 2-ketose.

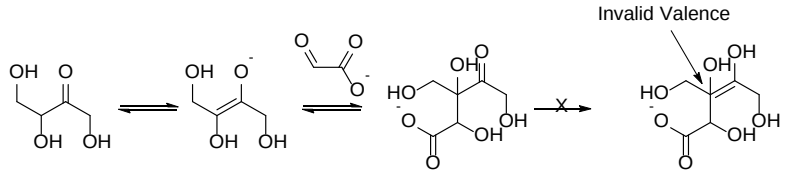


Figure 53 – Demonstration of the Impossibility of Carbonyl Migration in the Branched Product of Erythrulose and Glyoxylate.

Overall, reactions of glyoxylate with each of the ketoses and glyoxylate presented herein are consistent with the glyoxylose reaction mechanism presented in Figure 35. The sum of these reactions illustrates how a system rich in glyoxylate with a small amount of glycolaldehyde can lead to the production of longer sugar species.

3.2 High pH Conversion of Aldoses to Aldonic Acids by Reaction with DHF

In Section 3.1, I presented the glyoxylose mechanism for sugar formation in which glyoxylate functions in a manner analogous to the formose reaction. While this reaction does bypass the branched product formation endemic to the formose reaction, the issue of long term sugar stability is unresolved. One potential solution to this program can be found by returning to Eschenmoser's "glyoxylate scenario" [23]. Sagi, et al.[24] demonstrated that reactions of dihydroxyfumarate (DHF) in aqueous medium with formaldehyde, glycolaldehyde, glyceraldehyde, or glyoxylate result in ketoses and/or ulosonic acids (which could decarboxylate to form ketoses). This reaction of DHF with glyoxylate to form a ketose is not consistent with the glyoxoin reaction of DHF and glyoxylate at high pH to form tartrates (Chapter 2). The difference in products under these conditions indicates the presence of a pH "switch," likely related to the abundance of hydroxide, which can be exploited to control the product suite produced when DHF is reacted with glyoxylate and influence the reaction pathway either toward tartrate or dihydroxyacetone/pentulosonic acid (see Scheme 1 for generalized comparison). In this section, I demonstrate that this high pH "switch" is not unique to the DHF-glyoxylate system but is a general method for producing aldonic acids rather than ketoses when DHF is reacted with formaldehyde, glycolaldehyde, and glyceraldehyde. I also show this high pH "deoxalation" can be utilized in the conversion of 4, 5, and 6 carbon aldoses to 6, 7, and 8 carbon aldonic acids albeit at lower yields

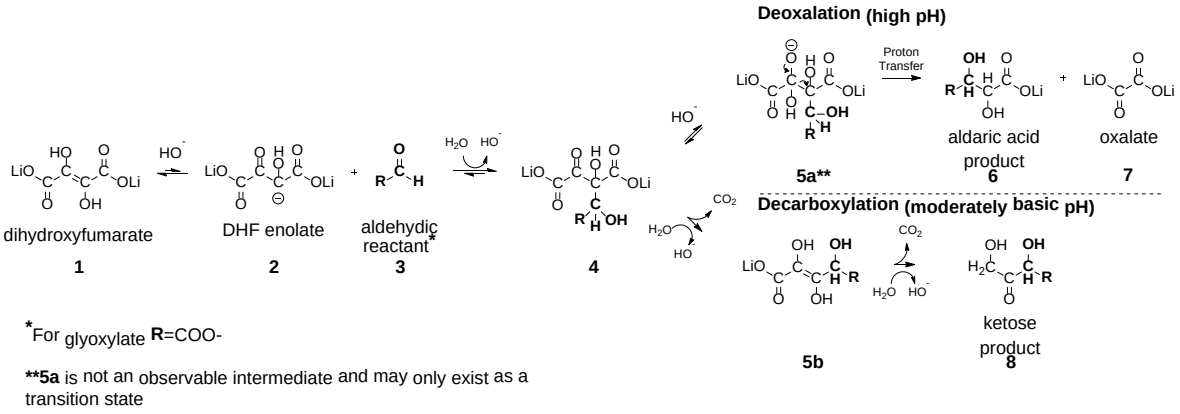


Figure 54 – Generalized Reaction of DHF and Aldehydes under Moderately and Highly Basic Conditions. Reaction of DHF **1** with an aldehyde **3** is initiated by deprotonation of DHF to form its enolate species **2**, which undergoes an aldol reaction with the free aldehyde, availability of which depends on the species equilibrium between monomer, hydrate, and oligomers. This aldol reaction produces a branched intermediate 4. It is at this point where the pH determines the ultimate products. Under high pH conditions the high concentration of available hydroxide results in a nucleophilic attack on the carbonyl of 4 by hydroxide, forming 5a which then fragments to form the 2-carbanion of the product acid, and bioxalate, which exchange a proton yielding an aldonic acid product 6 and oxalate 7. Under moderately basic conditions 4 undergoes a more straightforward decarboxylation reaction, yielding 5b which can undergo an enol to keto tautomerization to the three position which will result in another decarboxylation to a ketose.

3.2.1 Reaction of Dihydroxyfumarate with Paraformaldehyde

All our reactions and products were characterized extensively by ^{13}C NMR spectroscopy and LC-MS. Categorically, these high pH reactions occurred rapidly, with reaction of DHF with aldehydes taking place in minutes, and the competing decomposition of DHF taking place over several hours. While DHF is known to dimerize[25], this pathway was never observed under high pH conditions, instead mixtures of DHF at pH 14 exhibited only the decomposition to tartronate, formate, glycolate, and carbonate[26]. As we have already demonstrated the deoxalation reaction of DHF and glyoxylate to occur under a range of concentration and temperatures, and even to occur when either species is only present as an insoluble salt, our main aim in this study was to survey the versatility of this reaction[26]; therefore, we have elected to

conduct the majority of these reactions at high concentration, room temperature, and homogeneous conditions, which also rendered detection by 13C-NMR less onerous.

We began our investigations of the high pH reactivity of DHF and aldehydes with the simplest aldehyde, formaldehyde, which is expected to be a component of any prebiotic milieu[2-4]. Reflective of these conditions, 0.6M DHF and 0.6M paraformaldehyde (monomer basis[2]) were combined in 2.0M LiOH[3] at room temperature. Upon combination the reaction mixture heated noticeably while losing the characteristic purple brown of DHF and becoming clear with a faint yellow hue within minutes. The reaction was monitored by ^{13}C-NMR after 30 minutes. The NMR of the crude reaction mixture was extremely clean with four major peaks attributed to oxalate (173.15) and glycerate (180.01, 74.47, 65.22). Six additional minor peaks, comprising less than 5% of the total integral, correspond to formate, carbonate, glycolate, and tartronate, all of which are known breakdown products of dihydroxyfumarate. No signals corresponding to starting materials were observed (Figure 1).

The conversion of dihydroxyfumarate to glycerate takes place by way of an aldol reaction, followed by a hydroxide promoted fragmentation. In this pathway, DHF is deprotonated to its enolate form, which undergoes an aldol reaction with free

---

[2] In reality, this will yield slightly less than 0.6M formaldehyde monomer due to paraformaldehyde stoichiometry of H-$(OCH_2)_n$-OH → n HCHO + 1 H_2O, a fact which explains the minor presence of DHF decomposition products but no remaining formaldehyde.

[3] Lithium hydroxide is required as the base in this reaction due to low solubility of potassium and sodium salts of dihydroxyfumarate, and solubility of lithium hydroxide is the limiting factor in setting the reaction conditions. Lithium hydroxide reaches the limit of ready solubility at 2.0M; as two equivalents base are required to neutralize DHF and a third equivalent is required for fragmentation of the five carbon intermediate **4**, 0.6M in reactants represents the upper limit for batch reaction, although it is possible that this may be overcome in a semi batch or mixed counter ion system.

formaldehyde monomer, produced by dynamic equilibrium from paraformaldehyde[27]. After capturing a proton from solution, the resultant five carbon species is attacked at the keto- moiety by hydroxide, fragmenting in a deoxalation pathway yielding glycerate and oxalate in a 1:1 ratio (Scheme 1). This deoxalation pathway was validated in our previous work by reaction of DHF with ^{13}C dilabelled glyoxylate to form partially labelled tartrate[26]. Based on the stoichiometry of this reaction, and data obtained from LC-MS analysis, glycerate is formed in 75% crude yield with respect to DHF.

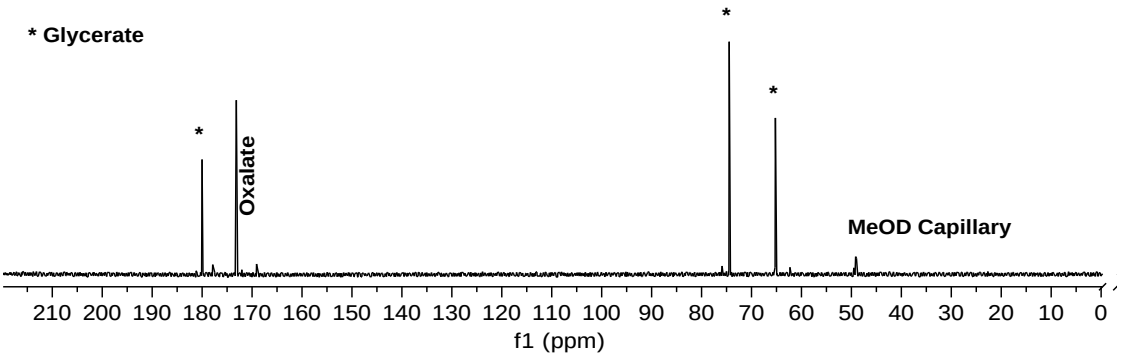


Figure 55 – NMR of Reaction of DHF and Paraformaldehyde After 30 minutes and comparison to standard. Reaction of 0.6M DHF and 0.6M paraformaldehyde in 2.0M LiOH after 30 minutes. No peaks associated with DHF or formaldehyde are observed. Peaks at 65.22, 74.47, and 180.01 ppm are consistent with glycerate. The signal at 173.15 ppm is the peak associated with oxalate. The remaining peaks are consistent with known DHF decomposition products. Lack of a peak corresponding to aqueous methanol suggests that DHF was in slight excess, and all formaldehyde was consumed by reaction with DHF.

In the work of Sagi, et al. DHF was observed to undergo a double addition of formaldehyde by way of an aldol reaction of DHF with formaldehyde followed by decarboxylation, and an additional aldol reaction with a second equivalent of formaldehyde. This branched acid species then underwent an additional decarboxylation, ultimately yielding tetrulose. Alternately, it is possible that this branched species could be formed and undergo a deoxalation reaction, or an analogous loss of glycolate (Figure 56).As a small amount of carbonate is evident in the reaction of one equivalent of DHF

with one equivalent of formaldehyde (Figure 55, 169ppm) there is some evidence this pathway may be occurring.

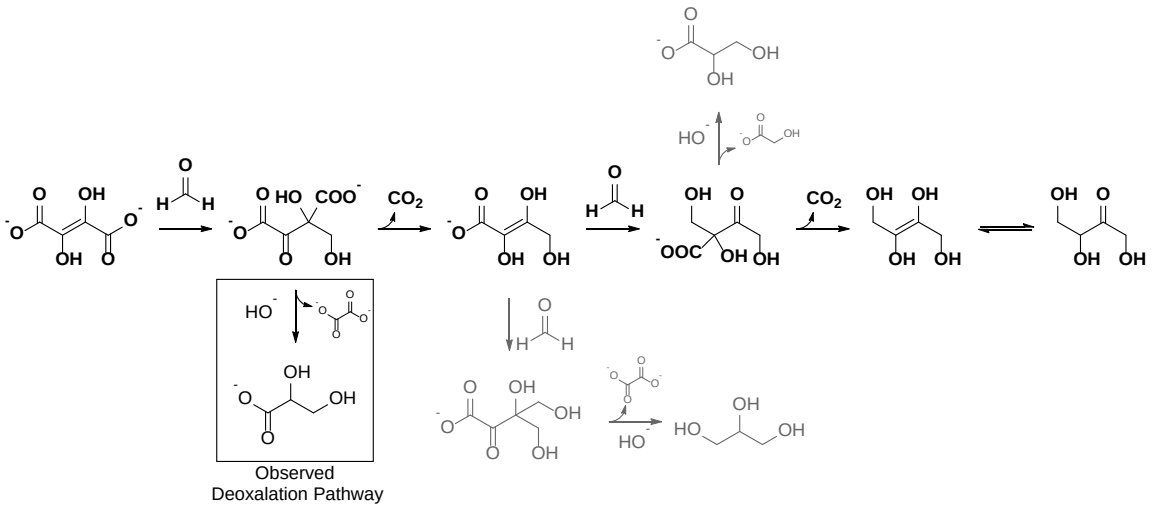


Figure 56 – Known and Hypothesized Reactions of DHF and Formaldehyde. Shown horizontally above is the mechanism proposed by Sagi et al. for the reaction of DHF with two equivalents of paraformaldehyde, yielding erythrulose. Shown in the box is the observed deoxalation reaction when DHF is reacted with one equivalent of paraformaldehyde at pH >13. Two potential side reactions which could occur when DHF is reacted with multiple equivalents of formaldehyde at high pH are shown in gray.

To investigate the reaction of DHF with multiple equivalents of formaldehyde we conducted a reaction of 0.6M DHF with 1.2M monomer equivalents of paraformaldehyde, as well as a reaction of 0.3M DHF with 0.6M monomer equivalents of paraformaldehyde. These reactions showed formation of by-products; however, the main product was glycerate in each case (Figure 57).

We also investigated the effect of low concentration and limited solubility by conducting reactions of 0.1 mmol of ^{13}C paraformaldehyde with the insoluble sodium salt of DHF in 20 and 100 mLs of 2.0M NaOH. In both cases a peak associated with glycerate labeled at the 3 position was observed.

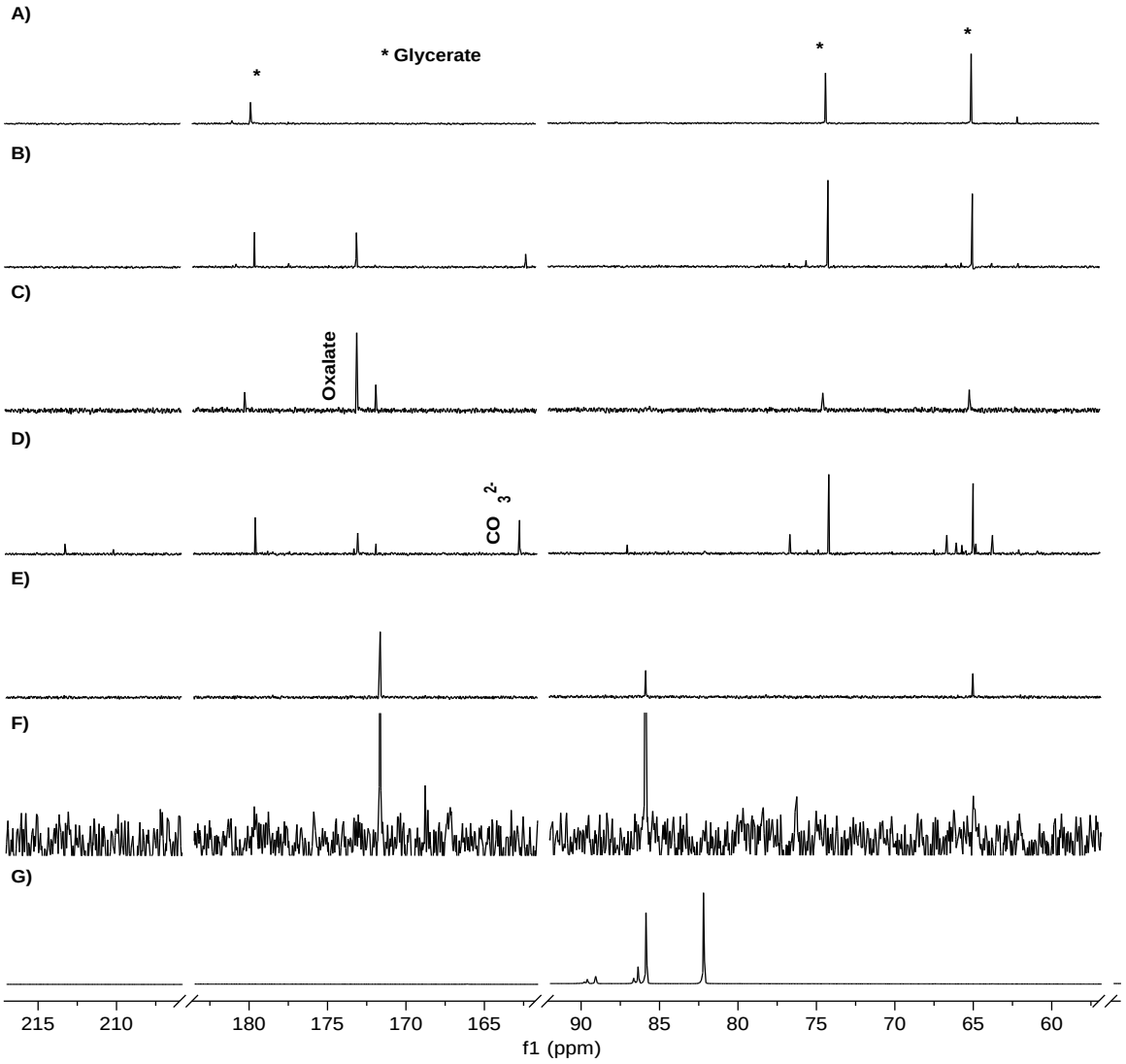


Figure 57 - Comparison of ^{13}C NMRs of the Reactions of DHF and Paraformaldehyde after 24 hours. A) Glyceric acid standard in 2.0M NaOH. B) Reaction of 0.6M DHF with 0.6M monomer equivalents paraformaldehyde. C) Reaction of 0.3M DHF with 0.6M monomer equivalents paraformaldehyde. D) Reaction of 0.6M DHF with 1.2M monomer equivalents paraformaldehyde. E) Reaction of .1mmol of ^{13}C paraformaldehyde (5mM) with 0.1M insoluble sodium DHF in 20mL 2M NaOH F) Reaction of .1mmol of ^{13}C paraformaldehyde(1mM) with 0.1M insoluble sodium DHF in 100mL 2M NaOH G) 0.6M monomer equivalents of paraformaldehyde in 2.0M LiOH showing distribution between monomer, dimer and higher oligomers. In B,C, and D, glyceric acid is the dominant product. In reaction D additional peaks in the 60-80 ppm range corresponding with CH-OH and CH2-OH peaks and the 205-215 ppm range corresponding to C=O peaks, suggest that a degree of formose reaction my be occurring under these conditions. Specifically this suggests that DHF reacts with a single equivalent of formaldehyde, then decarboxylates, preserving a keto moiety. This species can then undergo reaction with additional equivalents of formaldehyde in a formose type pathway. Even in this case however glycerate is clearly the major product.

3.2.2 Reaction of Dihydroxyfumarate with Glycolaldehyde

We next investigated the reaction of DHF with glycolaldehyde. 0.6M DHF and 0.3M glycolaldehyde dimer[4] were combined in 2.0M LiOH at room temperature. As was the case with the reaction of DHF and paraformaldehyde, the reaction heated noticeably while changing from dark to clear with a yellow tint. The reaction was monitored by ^{13}C-NMR after 1 hour, at which time unreacted DHF was still observed, and again after 24 hours. The NMR of the reaction mixture was more complex than in the case of formaldehyde, with a larger portion of DHF decomposition products. In addition to these decomposition products, 8 new peaks (179.91, 179.34, 74.69, 74.30, 73.79, 73.30, 63.82, 62.83ppm) were observed consistent, with the two diastereomers of the expected tetronic acid product. In neither the one hour nor the one day NMR were peaks consistent with glycolaldehyde observed.(Figure 2).

Comparison of the ^{13}C NMR of this reaction with the ^{13}C NMR of D-threono-1,4-lactone in 2M LiOH shows a one to one correlation of all peaks in the standard to the reaction products, with an additional set of off-set peaks which likely correspond to the erythro-diastereomer. Quantitation of the reaction yield was conducted using LCMS with the combined integral of threono- and erythro- diastereomers compared against a calibration curve prepared from the D-threono-1,4-lactone in 2M LiOH. This quantitation showed a crude yield of 70%.

---

[4] Resulting in a monomer concentration of 0.6M as shown by a standard in 2.0M LiOH. Monomerization was rapid, occurring in the less than 10 minutes between preparation of a standard in 2.0M LiOH and obtaining NMR data.

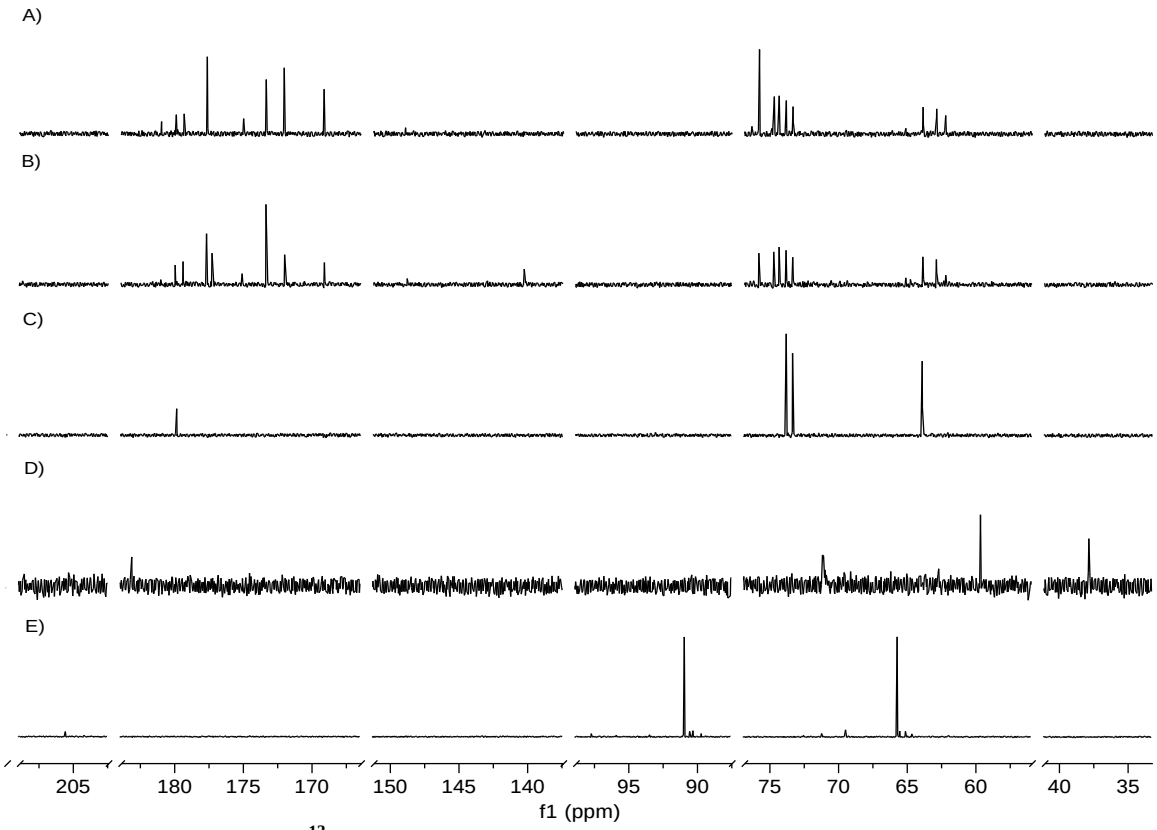


Figure 58 - Comparison of ^{13}C NMRs of the Reactions of DHF and Glycolaldehyde after 24 hours. Reaction of 0.6M DHF and 0.3M glycolaldehyde dimer (0.6M monomer after hydrolysis) in 2.0M LiOH after A) 1 hour and B) 1 day. Comparison to authentic threonic acid C) shows that peaks at 179.91, 73.79, 73.30, and 63.82 ppm are consistent with this species. Peaks at 179.34, 74.69, 74.30, and 62.83ppm are assumed to be the diasteromer, erythronic acid. Comparison to glycolaldehyde d) shows that no peaks associated with glycolaldehyde are observed in either spectrum, and DHF (139 ppm) is observed only in the 1 hour data set. The signal at 173.15 ppm is the peak associated with oxalate. The remaining peaks large peaks are consistent with known DHF decomposition products, however some small peaks in the 60-80 ppm range are indicative of side product formation. Comparison to the reaction of 0.6M glycolaldehyde in 2.0M LiOH after one day, shows that self reaction of glycolaldehyde is not important to the reaction pathway.

3.2.3 Reaction of Dihydroxyfumarate with Glyceraldehyde

The reaction of 0.3M glyceraldehyde dimer and 0.6M DHF was conducted in 2.0M LiOH, yielding the expected pentonic acid product, confirmed by ^{13}C-NMR and LCMS. The pentonic acid product of the reaction of DHF and glyceraldehyde has four diastereomeric products, yielding twelve peaks in the 70-75 ppm range. These can be seen and compared to the spectrum of ribonic acid, which shows good correlation between reaction outcome and standard(Figure 3). LCMS data shows the expected mass

to charge ratio of 165 in negative ion mode as the only major peak of this reaction after only 10 minutes, with the peak growing in intensity after 2 hours. Quantitation of the reaction yield was conducted using LCMS with the combined integral of the diastereomers compared against a calibration curve prepared from the gamma lactone of ribonic acid in 2M LiOH. This quantitation showed a yield of 8.5% of pentonic acid.

Sagi, et al. showed DHF to selectively react with aldoses under conditions of pH 8-9 by comparing reactivity of DHF with glyceraldehyde and with the analogous ketose, dihydroxyacetone(DHA). We sought to test whether this selectivity continued under pH 14 conditions. Accordingly, the reaction of 0.6M DHF and 0.3M DHA dimer in 2.0M LiOH was conducted with the expectation of no reaction occurring between DHF and DHA. Surprisingly, ^{13}C-NMR of this reaction showed a moderate peak at 173 ppm consistent with production of oxalate, indicating reaction with DHF is occurring, although seemingly to a lesser extent than the reactions discussed previously. Comparison of the ^{13}C NMR of this reaction with that of the glyceraldehyde reaction reveals no similarities besides the oxalate peaks and peaks corresponding to decomposition of DHF, and shows the DHA spectrum to have a greater number of smaller peaks. LCMS analysis of this reaction showed peaks corresponding to oxalate, hexose, and octonic acid with no peak corresponding to a hexonic acid product. Theses masses seem to indicate that DHA isomerized to, and condensed with glyceraldehyde to form a ketohexose[28] which either reacted with DHF directly through the deoxalation pathway, or tautomerized to an aldohexose[29] and subsequently reacted with DHF(Scheme 2).

As the chain length of the expected aldonic acid product increases, ^{13}C-NMR rapidly loses utility for product identification due to the number of diastereomers produced. Consequently we rely more heavily on mass spectrometry to determine outcome of reaction of these larger sugars.

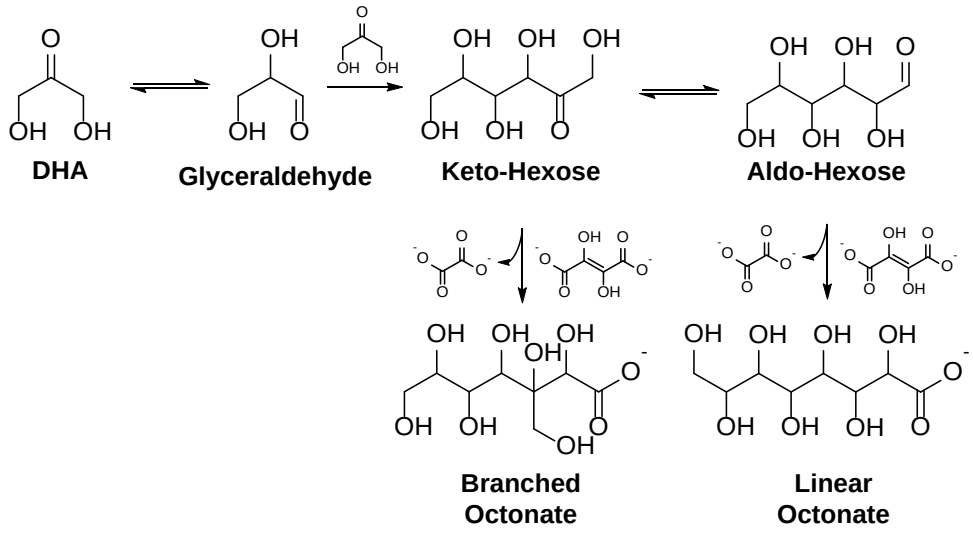


Figure 59 – Proposed Scheme of Self Condensation of Dihydroxyacetone and Subsequent Reaction with Dihydroxyfumarate to Form Octonate. The condensation of trioses to form hexose is a well known reaction[30]. Based on observation of a mass spectral signal of 255 m/z in the reaction of 0.6M DHA with 0.6M DHF. This peak was not observed in the analogous glyceraldehyde reaction, suggesting these may be branched products.

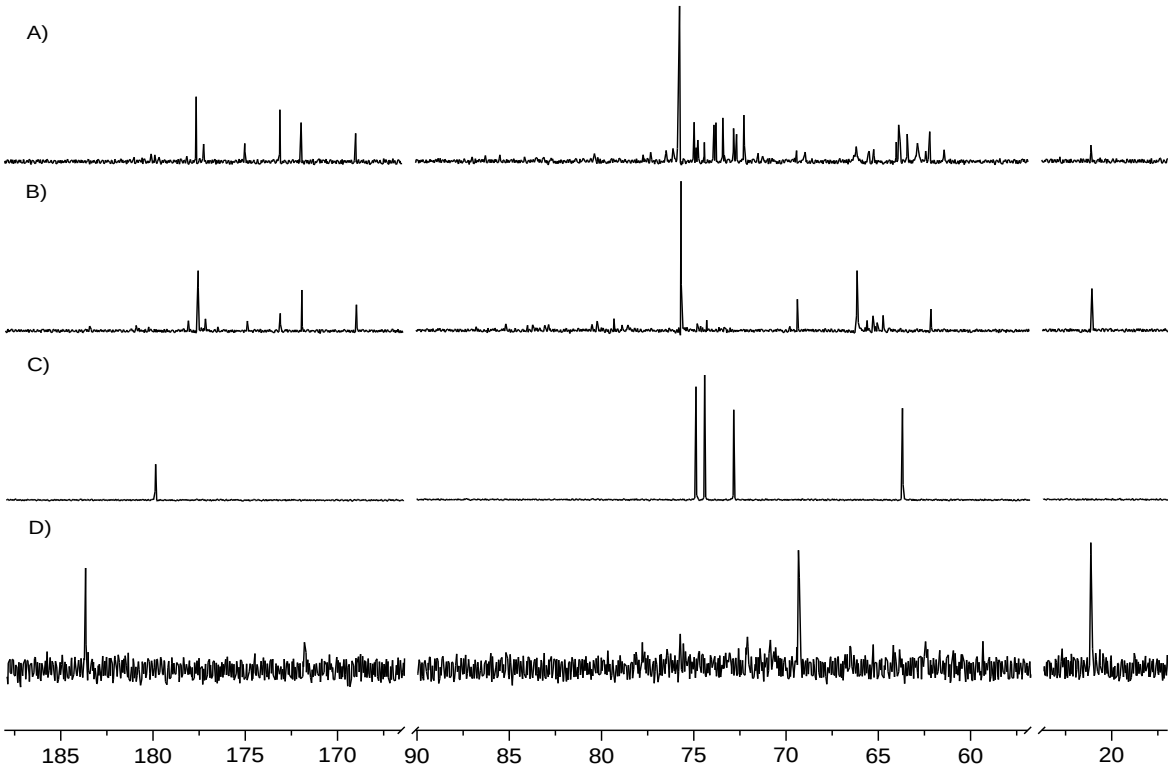


Figure 60 - Comparison of ^{13}C NMRs of the Reactions of DHF and Glyceraldehyde/DHA after 24 hours. A) Reaction of 0.6M glyceraldehyde with 0.6M DHF in 2.0M LiOH B) Reaction of 0.6M DHA with 0.6M DHF in 2.0M LiOH. C) 0.6M ribonic acid gamma lactone in 2.0M NaOH. D) Control reaction of 0.3M glyceraldehyde in 2.0M LiOH. Comparison of these reactions shows a 1:1 correlation between sodium ribonate with product peaks in A). Ribonate is one of 4 diastereomers of pentonic acid; there appear to be 12 peaks in the 70-75 ppm range and more than 4 peaks in the 60-65 ppm range of A) suggesting that the other 3 diastereomers are formed as well. Some small peaks in B) suggest production of a larger species which we suggest may be isomers of octonate based on mass spectral evidence. The self condensation of glyceraldehyde does not seem to be a large component of either A) or B).

Table 5 - Yields of Reaction of DHF with Aldehydes and Aldoses with Analogous Keto Species for Comparison

DHF was reacted with various aldehydes and ketones at a concentration of 0.3M in each reactant at room temperature in a solution of aqueous 2.0M LiOH. After 24 hours reaction mixtures were diluted at a 10000:1 ratio and analyzed by LCMS.

Aldehydic Reactant	Aldonic Acid Yield (%)	Mole Fraction of Free Aldehyde	Reference	Keto Reactant	Yield(%)
Formaldehyde	75	0.03	Van Wazer[27]	---	---
Glycolaldehyde	8.5	0.04	Yaylayan[31]	---	---
Glyceraldehyde	13.5	0.04		---	---
Glyceraldehyde(dimer)	27.8	0.044	Wolfenden[32]	Dihydroxyacetone(dimer)	---
Erythrose	32.5	0.0094*	Seriani[33]	Erythrulose	28.3
Ribose	5.6	0.00050	Miller[34]	Ribulose	0.43
Xylose	4.7	0.00020	Miller[34]	---	---
Glucose	1.09	2.51-05	Miller[34]	Fructose	0.63
Mannose	1.8	5.01-05	Miller[34]	---	---
Glyoxylic Acid	82	0.00333	Sorensen [35]		

*Threose Value

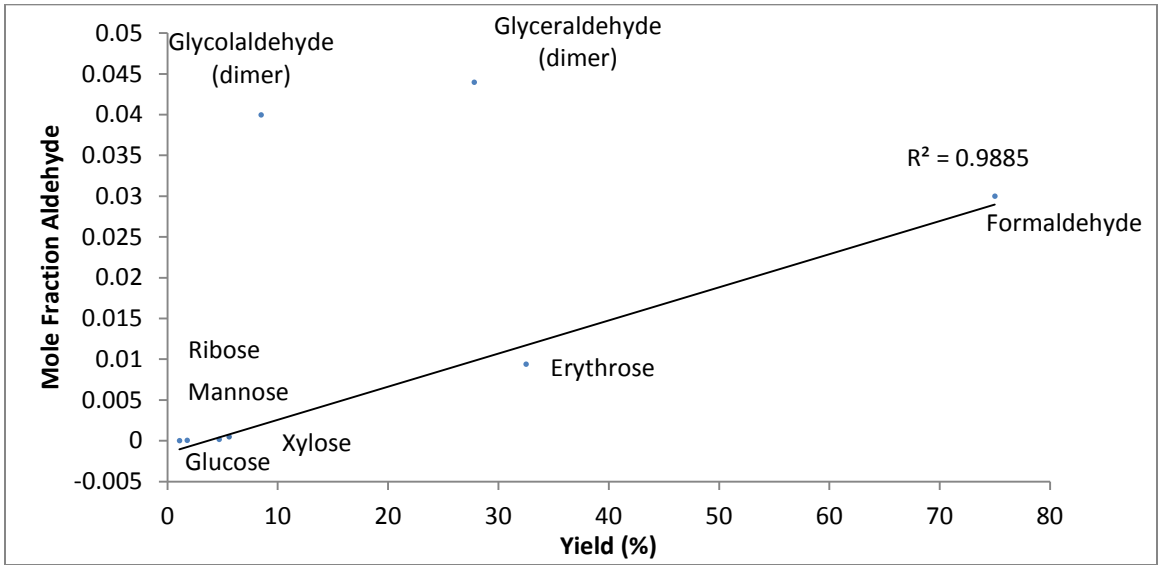


Figure 61 - Yields of Reaction of DHF with Aldehydes and Aldoses vs Free Aldehyde Concentration

3.2.4 Reaction of Dihydroxyfumarate with Higher Aldoses

DHF was also reacted with several aldoses to test the propensity for the "deoxalation" reaction to occur with these higher sugars. 0.6M DHF was reacted with 0.6M of each sugar in 2.0M LiOH, and the result quantified by LCMS the results of which can be seen in Table 1. Calibration curves were prepared from gluconate and glucoheptonate in 2.0M LiOH with the glucoheptonate standard being used for both the 7 and 8 carbon products. The yield of the aldose reactions decreased with increasing sugar chain length, a trend that is true across all the aldehyde substrates. In fact, the yield of the deoxalation reaction correlates well to the concentration of free aldehyde of the substrate(with the exception of glycolaldehyde and glyceraldehyde, possibly due to their initial dimer form), a fact which may be useful in predicting reactivity of other molecules under similar conditions (Figure 61). This trend is highlighted specifically by the increased yield of mannose compared to glucose, and ribose compared to xylose, where both ribose and mannose have a higher proportion of free aldehyde.

Based on this observation of greater reactivity of species with higher proportion of free aldehyde, we attempted to improve the reaction yield by adding cyanide to the reactions of DHF and ribose and DHF and glucose. Cyanide is known to react with aldoses to form aldonitriles which will hydrolyze under basic conditions, also to aldonic acids but with one additional carbon [36] rather than the two additional carbons afforded by reaction with DHF. Our experience with the reaction of cyanide and glyoxylate suggests that the existence of the cyanohydrin form of glyoxylate does not interfere with quantitative reaction with DHF; hydrolysis of the cyanohydrin seems to be considerably slower than reverse equilibrium to the aldehyde and reaction via the deoxalation pathway. When 1 equivalent of cyanide were added to 0.5M glucose and 0.5M DHF the reaction

increased from 1% yield to 34%. Ribose, which began with a 5% yield of the aldonic acid species, increased to a 35% yield in the presence of one equivalent of cyanide.

Additionally, DHF was reacted with the keto forms of some of the aldoses investigated above, again with the expectation of no deoxalation product forming. These ketoses revealed products with m/z consistent with the aldose products (Table 5). Further experiments will be carried out to identify whether branched products are formed, or if the aldonic acids are formed from isomerization from the keto to aldose form.

3.2.5 Interception of DHF Produced in the Glyoxoin Reaction

In Chapter 2, I indirectly demonstrated the formation of DHF under high pH conditions by the cyanide mediated dimerization of glyoxylate. These results suggested that DHF is produced from glyoxylate and cyanide, but is rapidly consumed by reaction with an additional equivalent of glyoxylate to form tartrate. The work in Section 3.2 was motivated by a desire to unify the previously demonstrated chemistry of the glyoxylate scenario[24], with these high pH conditions necessary for DHF formation [26]. Thus far, the results presented herein have assumed availability of DHF, but I have yet to show this availability. Section 3.2 was also motivated by a desire to unify the previously demonstrated chemistry of the glyoxylate scenario[24] , with the high pH conditions necessary for DHF formation[26]. To test whether the (presumed) DHF produced in the glyoxoin reaction[26] would be available to reactants other than glyoxylate, we conducted a reaction of 1.0M glyoxylic acid, 0.5M (monomer equivalent) paraformaldehyde, and 0.1M NaCN. This reaction was monitored by ^{13}C NMR after 1 day. Peaks consistent with glycerate, the expected aldonic acid product of DHF and

formaldehyde were observed (Figure 62). By mass spectroscopy this reaction resulted in a 25% yield of glycerate with respect to formaldehyde.

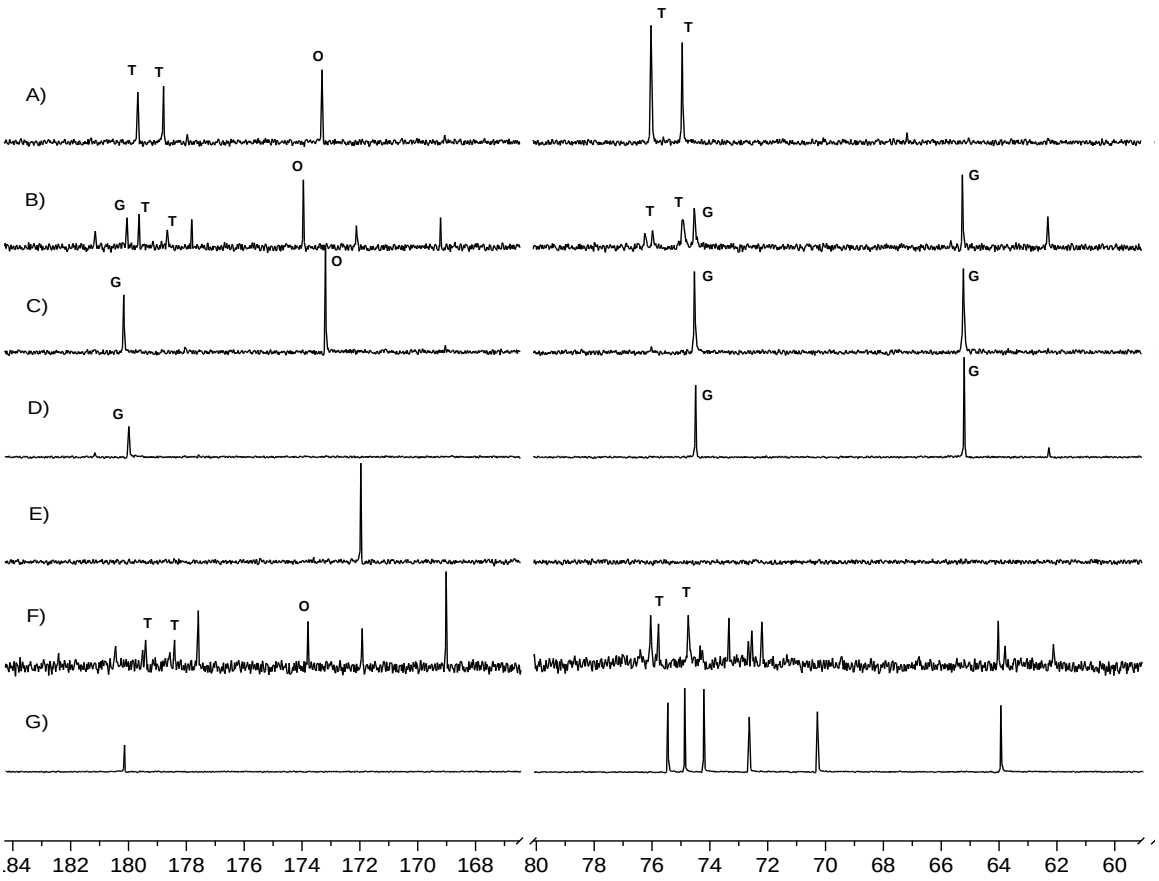


Figure 62 - ^{13}C NMR Data Demonstrating the Availability of Dihydroxyfumarate Produced by the Glyoxoin Reaction for Reaction with Aldehydes Other Than Glyoxylate. A) Reaction of 1.0M glyoxylic acid and 0.1M NaCN in 2.0M NaOH after 1 day. B) Reaction of 1.0M glyoxylic acid, 0.5M monomer equivalents of paraformaldehyde and 0.1M NaCN in 2.0M NaOH after 1 day. C) Reaction of 0.5M DHF and 0.5M monomer equivalents of paraformaldehyde in 2.0M LiOH after 1 day. D) 0.5M Glyeric acid standard in 2.0M NaOH E) Control reaction of 0.5M monomer equivalents of paraformaldehyde with 0.1M NaCN after 1 day. F) Reaction of 1.0M glyoxylic acid, 0.5M erythrose and 0.1M NaCN in 2.0M NaOH after 1 day. G) 0.5M Sodium gluconate standard in 2.0M NaOH. Conducting the glyoxoin reaction in the presence of paraformaldehyde B) clearly results in a mixture of products corresponding to tartrates (**T**) from reaction of DHF and glyoxylate A), and glycerate (**G**) from reaction of DHF and formaldehyde C) as well as the expected oxalate (**O**) byproduct. Conducting the glyoxoin reaction in the presence of erythrose F) demonstrates additional product peaks in the 70-80ppm range which correspond fairly well with the peaks associated with gluconate G) suggesting production of another diasteromer of hexose.

We also decided to test whether produced DHF would be available to react with higher molecular weight aldoses, which we have observed to be significantly less reactive

than formaldehyde. To this end, we conducted a reaction of 1.0M glyoxylic acid, 0.5M erythrose, and 0.1M NaCN. This reaction was monitored by ^{13}C NMR after 1 day. Peaks consistent with a hexose diasteromer are visible in the 13C NMR spectrum, and mass spectrometry data shows the production of hexose with a 13% yield with respect to erythrose. Combined with the result from the glyoxoin reaction in the presence of paraformaldehyde, these results show that while the reaction of DHF at high pH is rapid the DHF produced by the glyoxoin reaction can be intercepted and diverted into the production of aldonic acids.

3.2.6 Summary

Prebiotically, the production of aldonic acids is interesting for a number of reasons. The most obvious potential prebiotic role for these species is as a more stable "stockpile" form for standard sugars, a possibility considered by Larralde, et al.,[34] when considering the suitability of ribose as a prebiotic polymer backbone material. They concluded that the low stability of ribose necessitates storage in a more stable form, suggesting that reaction with cyanide through a Kiliani reaction to form an aldonic acid could provide such storage. These aldonic acids could then undergo an oxidation event[5] at the three position which would cause a decarboxylation yielding a ketose, which can tautomerize to an aldose. Under Kiliani conditions this process of aldose→aldonic acid→ketose→aldose results in reformation of an aldose with the same carbon length as the original. By utilizing DHF for the initial aldose to aldonic acid conversion, this

[5] Larralde, et al. suggest an early enzyme to be responsible, however mineral oxidations could also contribute [12]

process can be coupled with chain growth with the final aldose having one additional carbon.

A second potentially impactful result of this chemistry is the ready production of glyceric acid from DHF and formaldehyde or from glyoxylate, cyanide, and formaldehyde. Orgel and Kolb[37] have shown that glyceric acid can be phosphorylated by combination with trimetaphosphate in alkaline media. Phosphorylated glyceric acid has an important role in modern glycolysis as a building block of ribose in the Pentose Phosphate Pathway [38-40] and as a phosphate transfer agent in lower glycolysis[41-43] . Proscribing to the theory that glycolysis was one of the earliest of metabolic pathways[44] leads to the conclusion that glyceric acid, once phosphorylated, could supply a similar function to the prebiotic metabolism. Even without being phosphorylated, glyceric acid has the potential to drive the prebiotic metabolism through formation of high energy esters[45], the breakdown of which could be coupled to other metabolic processes[46]. These high energy esters could be formed through evaporative dehydration, a process which is known to effect the polyesterification of lactic acid[47], a reasonable structural analog to glyceric acid. It has also been shown that these polyglycerates can be formed catalytically by the presence of the ethanethiol-thioester of glyceric acid[48], and by simple heating in aqueous solution, although this route can be catalyzed by sulfuric acid[49].

Oxidized forms of these sugar acids could also be used as a prebiotic source of early structural materials. While modern biology primarily uses non-carboxylated sugars, sugar acids are incorporated into a number of structural polysaccharides. Hexuronic acids are incorporated into glycosaminoglycans[50] and pectins[51], important structural

polysaccharides in both animals and plants, respectively, to enhance water retention. Hexuronic acids are also incorporated into bacterial secretions to create hydrogels such as xanthan gum[52] and mucins[53]. These polymers are secreted by bacteria to help control their local environments by retaining water, ions and useful chemicals. It is not unreasonable to believe that early sugar acids could create a similar gel which could have favorable properties for collecting the components of early life [54-56].

Combined, the ability of glyoxylate to produce free sugars through the glyoxylose reaction as well as sugar acids by high pH reaction of DHF (formed from glyoxylate dimerization) with aldoses (obtainable by isomerization of ketoses produced in the glyoxylose reaction) shows how glyoxylate can function under various conditions as Hadean source molecule for carbohydrate production. The ability for DHF to react with paraformaldehyde to produce glycerate, which has important metabolic roles, combined with the ability of DHF and glyoxylate to lead to intermediates of the citric acid cycle, suggest a potentially important role in the origin of metabolism. Nor are these pathways mutually exclusive, pH gradients at hydrothermal vents have been proposed as a means of driving early metabolism[57] and such a system at a very high pH vent combined with a source of glyoxylate, could allow both high and moderate pH processes to occur in close proximity where products of each may be able to diffuse and interact. Overall, this underscores the need to identify a prebiotically plausible route to glyoxylate formation.

3.3 Experimental

Materials: Lithium hydroxide, sodium cyanide, glyoxylic acid monohydrate, dihydroxyfumaric acid dehydrate, glyceric acid, D-threono-1,4-lactone, ribonic acid gamma lactone, sodium gluconate, D-glucoheptono-1,4-lactone, paraformaldehyde were all purchased from Sigma Aldrich. Sodium hydroxide was obtained from EMD Millipore. Di-labeled ^{13}C glyoxylic acid (CLM-6027-PK; Lot PR-11812), and singly labeled ^{13}C sodium cyanide (CLM-378-1; Lot PR-20741) both with 99% isotopic substitution were purchased from Cambridge Isotope Laboratories Inc.

NMR: ^{13}C NMR data were obtained using either a Bruker Avance II-500 spectrometer or a Bruker Avance III-400 spectrometer. All NMR data were obtained using a 5mm probe, with a 1mm d_4-MeOH capillary as external standard and lock solvent. Data were collected using an inverse gated, 30 degree, ^{13}C pulse sequence with a D1 of 45 seconds.

Quantitative NMR results were assured mathematically, using the methods given by Traficante [58]. For this analysis, product T_1 measurements were obtained from an inversion recovery experiment. The longest T_1's were those of oxalate and formate at 32 seconds. Based on this T_1 and a 30 degree pulse, we calculate a 95% relaxation of these species with a delay time of 45 seconds. At 300 seconds this relaxation is 99.998%. This indicates that an error of no larger than 5% will be observed with a delay time of 45 seconds. This mathematical result was also tested experimentally as shown in Figure 31. These NMR conditions allow us to collect reasonable quantitative data in 30% of the time required under the conditions of a 90 degree pulse and $5T_1$ relaxation time.

pH: pH greater than 10 was measured using Aldrich Whatman brand PANPEHA pH Paper (Sensitivity of 0.5 pH unit in the 0 to 9 range; 1 pH unit in the 9 to 14 range). At high pH, measurements using glass electrodes become inaccurate; therefore, pH paper was used. At pH less than 10 a Cole Palmer pH 20 system was utilized.

Mass Spectrometry: Mass spectrometry data were collected using a Waters 3100 Mass Detector without chromatographic separation. The electrospray ionization injector was operated in both positive and negative mode. The mobile phase was a 1:1 water/AcN mixture with 5mM ammonium formate.

MRM monitoring: Multiple reaction monitoring MS/MS data were obtained using a Shimadzu LCMS MS-8038 system operating without a column in a pure water mobile phase.

Stock Solutions: Stock solutions of LiOH and NaOH were prepared by weighing the required mass of the hydroxide into a browned glass bottle. The necessary amount of LCMS grade H_2O was added as measured by graduated cylinder. The bottle was then closed and subjected to sonic mixing until complete dissolution was achieved. Neither degassing nor inert atmosphere was required.

Unbuffered Glyoxylose Reactions: Reaction was conducted at room temperature by combining known masses of reactants in LCMS grade water and adjusting the pH using aqueous NaOH.

Buffered Glyoxylose Reactions: Reaction was conducted at room temperature by adding a known masses of glyoxylic acid monohydrate to a phosphate buffer along with an equal number of moles of sodium hydroxide to preserve buffer pH. When the reaction cooled a desired amount of labeled or unlabelled carbohydrate reactant was added to the solution. When carbohydrates were provided in the form of a stock solution (e.g. 2-^{13}C glycolaldehyde) the initial phosphate buffer was prepared at twice the desired final concentration, and sufficient LCMS water was added such that addition of the reactant stock solution would bring the total volume to twice the initial.

Derivatization with 2-nitrophenylhydrazine: 2-Nitrophenyl hydrazine (solid, 0.1847g, 0.0012 mol, 0.24 M) was measured into a large test tube with stir bar to which 4 mL LCMS EtOH was added. Add 1 mL reaction solution to test tube. Heat derivatization solution to a low boil for 20 min. Cool and dilute for LCMS analysis. LC Separation was conducted using the following method on a C18 Column at a flow rate of 0.3 mL/min.

Time	AcN (%)	H_2O 1 mM NH_4 HCOO pH 3.5 (%)
0	5	95
3	5	95
18	95	5
21	95	5
25	5	95

Reaction of Glyoxylate, Cyanide, and Paraformaldehyde/Erythrose: Glyoxylic acid monohydrate and sodium cyanide and either erythrose or paraformaldehyde were each weighed in individual vials. The necessary volume of 2.0 M NaOH or 2.0 M LiOH were

then added to the vial containing glyoxylate. This vial was mixed until no solid remained. The solution was then transferred by syringe to the vial containing either erythrose or paraformaldehyde. This second vial was mixed until no solid remained. The solution was then transferred by syringe from the second vial to the vial containing cyanide. This third vial was mixed until no solid remained at which time the reaction was allowed to proceed. The pH measurements of the resulting solutions were > 13.

Reaction of Dihydroxyfumarate with Paraformaldehyde or Carbohydrates:

Dihydroxyfumaric acid dehydrate, and the desired co-reactant were each weighed in individual vials. LiOH or NaOH was added to the vial containing the co-reactant. This vial was mixed until no solid remained. The solution was then added to the vial containing the dihydroxyfumaric acid. The pH of the resulting solution was > 14.

3.4 References

[1] BRESLOW R. On the mechanism of the formose reaction. Tetrahedron Lett 1959;1:22-6.

[2] MILLER SL. A Production of Amino Acids under Possible Primitive Earth Conditions. Science 1953;117:528-9.

[3] CLEAVES HJ. The prebiotic geochemistry of formaldehyde. Precambrian Research 2008;164:111-8.

[4] LOB W. Uber das Verhaltes des formamibs unter ber wirkung derstillen ent-adung: ein beitrag zue frage der flictoff-assimilton'. Chem Ber 1913;46:690.

[5] DECKER P, et al. Bioids: X. Identification of formose sugars, presumable prebiotic metabolites, using capillary gas chromatography/gas chromatography—mas spectrometry of n-butoxime trifluoroacetates on OV-225. Journal of Chromatography A 1982;244:281-91.

[6] YOSHIMURA J. Synthesis of branched-chain sugars. Advances in carbohydrate chemistry and biochemistry 1984;42:69-134.

[7] KIM H-J, et al. Synthesis of carbohydrates in mineral-guided prebiotic cycles. J Am Chem Soc 2011;133:9457-68.

[8] BENNER SA, et al. Asphalt, water, and the prebiotic synthesis of ribose, ribonucleosides, and RNA. Accounts of chemical research 2012;45:2025-34.

[9] SPECK JR JC. The lobry de Bruyn-Alberda van Ekenstein transformation. Advances in carbohydrate chemistry 1958;13:63-103.

[10] SCHWARTZ A, DE GRAAF RM. The prebiotic synthesis of carbohydrates: A reassessment. Journal of Molecular Evolution 1993;36:101-6.

[11] VARMA R, FRENCH D. Mechanism of the cyanohydrin (Kiliani-Fischer) synthesis. Carbohydrate Research 1972;25:71-9.

[12] NAIDJA A, SIFFERT B. Oxidative decarboxylation of isocitric acid in presence of montmorillonite. Clay Minerals 1990;25:27-37.

[13] HORECKER BL. The pentose phosphate pathway. Journal of Biological Chemistry 2002;277:47965-71.

[14] MICHAEL EGGERT F, JONES M. Measurement of neutral sugars in glycoproteins as dansyl derivatives by automated high-performance liquid chromatography. Journal of Chromatography A 1985;333:123-31.

[15] BRESLOW R, APPAYEE C. Transketolase reaction under credible prebiotic conditions. Proceedings of the National Academy of Sciences 2013;110:4184-7.

[16] KIM J-S, LEE Y-S. Effect of reaction pH on enolization and racemization reactions of glucose and fructose on heating with amino acid enantiomers and formation of melanoidins as result of the Maillard reaction. Food Chemistry 2008;108:582-92.

[17] VENKATARAMAN R, RACKER E. Mechanism of action of transaldolase I. Crystallization and properties of yeast enzyme. Journal of Biological Chemistry 1961;236:1876-82.

[18] SROKOL Z, et al. Hydrothermal upgrading of biomass to biofuel; studies on some monosaccharide model compounds. Carbohydrate research 2004;339:1717-26.

[19] HARMAN CE, et al. Atmospheric Production of Glycolaldehyde Under Hazy Prebiotic Conditions. Origins of Life and Evolution of Biospheres 2013;43:77-98.

[20] HOLLIS JM, et al. Interstellar glycolaldehyde: the first sugar. The Astrophysical Journal Letters 2000;540:L107.

[21] BENNETT CJ, KAISER RI. On the formation of glycolaldehyde (HCOCH2OH) and methyl formate (HCOOCH3) in interstellar ice analogs. The Astrophysical Journal 2007;661:899.

[22] SEEHOLZER SH, et al. Enolpyruvate: chemical determination as a pyruvate kinase intermediate. Biochemistry 1991;30:727-32.

[23] ESCHENMOSER A. The search for the chemistry of life's origin. Tetrahedron 2007;63:12821-43.

[24] SAGI VN, et al. Exploratory Experiments on the Chemistry of the “Glyoxylate Scenario”: Formation of Ketosugars from Dihydroxyfumarate. J Am Chem Soc 2012;134:3577-89.

[25] NAIDU SAGI V, et al. Diastereoselective Self-Condensation of Dihydroxyfumaric Acid in Water: Potential Route to Sugars. Angewandte Chemie International Edition 2011;50:8127-30.

[26] BUTCH C, et al. Production of Tartrates by Cyanide-Mediated Dimerization of Glyoxylate: A Potential Abiotic Pathway to the Citric Acid Cycle. J Am Chem Soc 2013;135:13440-5.

[27] MOEDRITZER K, WAZER JRV. Equilibria between cyclic and linear molecules in aqueous formaldehyde. The Journal of Physical Chemistry 1966;70:2025-9.

[28] GUTSCHE CD, et al. Base-Catalyzed Triose Condensations. J Am Chem Soc 1967;89:1235-45.

[29] SOWDEN JC, SCHAFFER R. The Reaction of D-Glucose, D-Mannose and D-Fructose in 0.035 N Sodium Hydroxide at 35°. J Am Chem Soc 1952;74:499-504.

[30] BERL WG, FEAZEL CE. The Kinetics of Hexose Formation from Trioses in Alkaline Solution1. J Am Chem Soc 1951;73:2054-7.

[31] YAYLAYAN VA, et al. Investigation of the mechanism of dissociation of glycolaldehyde dimer (2, 5-dihydroxy-1, 4-dioxane) by FTIR spectroscopy. Carbohydrate research 1998;309:31-8.

[32] WOLFENDEN RV. Binding of substrate and transition state analogs to triosephosphate isomerase. Biochemistry 1970;9:3404-7.

[33] SERIANNI AS, et al. Anomerization of furanose sugars: kinetics of ring-opening reactions by proton and carbon-13 saturation-transfer NMR spectroscopy. J Am Chem Soc 1982;104:4037-44.

[34] LARRALDE R, et al. Rates of decomposition of ribose and other sugars: implications for chemical evolution. Proceedings of the National Academy of Sciences 1995;92:8158-60.

[35] SORENSEN PE, BRUHN, K., LINDELOV, F. Kinetics and equilibria for the reversible hydration of the aldehyde group in glyoxylic acid. Acta Chem Scand A 1974;28.

[36] SERIANNI AS, et al. Cyanohydrin synthesis: studies with carbon-13-labeled cyanide. The Journal of Organic Chemistry 1980;45:3329-41.

[37] KOLB V, ORGEL L. Phosphorylation of glyceric acid in aqueous solution using trimetaphosphate. Orig Life Evol Biosph 1996;26:7-13.

[38] WOOD HG. Significance of alternate pathways in the metabolism of glucose. Physiol Rev 1955;35:841.

[39] CAMPBELL AE, et al. Reductive pentose phosphate cycle in Nitrosocystis oceanus. Journal of bacteriology 1966;91:1178-85.

[40] GUNSALUS I, et al. Pathways of carbohydrate metabolism in microorganisms. Bacteriological reviews 1955;19:79.

[41] ROSE IA, WARMS JV. Control of glycolysis in the human red blood cell. Journal of Biological Chemistry 1966;241:4848-54.

[42] HARKNESS DR, et al. A comparative study on the phosphoglyceric acid cycle in mammalian erythrocytes. Comparative biochemistry and physiology 1969;28:129-38.

[43] LIPMANN F. Metabolic generation and utilization of phosphate bond energy. Adv Enzymol Relat Areas Mol Biol 1941;1:99-162.

[44] ALBERTS B. Molecular Biology of the Cell: Reference edition: Garland Science; 2008.

[45] WEBER A, HSU V. Energy-rich glyceric acid oxygen esters: Implications for the origin of glycolysis. Orig Life Evol Biosph 1990;20:145-50.

[46] WEBER A. The triose model: Glyceraldehyde as a source of energy and monomers for prebiotic condensation reactions. Orig Life Evol Biosph 1987;17:107-19.

[47] MAMAJANOV I, et al. Ester Formation and Hydrolysis during Wet–Dry Cycles: Generation of Far-from-Equilibrium Polymers in a Model Prebiotic Reaction. Macromolecules 2014;47:1334-43.

[48] WEBER AL. Oligoglyceric acid synthesis by autocondensation of glyceroyl thioester. Journal of molecular evolution 1987;25:191-6.

[49] WEBER AL. Thermal synthesis and hydrolysis of polyglyceric acid. Orig Life Evol Biosph 1989;19:7-19.

[50] MEYER K. The biological significance of hyaluronic acid and hyaluronidase. Physiol Rev 1947;27:19-47.

[51] WILLATS WG, et al. Pectin: cell biology and prospects for functional analysis. Plant Cell Walls: Springer; 2001. p. 9-27.

[52] BECKER A, et al. Xanthan gum biosynthesis and application: a biochemical/genetic perspective. Applied microbiology and biotechnology 1998;50:145-52.

[53] BANSIL R, TURNER BS. Mucin structure, aggregation, physiological functions and biomedical applications. Current Opinion in Colloid & Interface Science 2006;11:164-70.

[54] TOLSTOGUZOV V. Why are polysaccharides necessary? Food Hydrocolloids 2004;18:873-7.

[55] TREVORS JT, POLLACK GH. Hypothesis: the origin of life in a hydrogel environment. Progress in Biophysics and Molecular Biology 2005;89:1-8.

[56] TREVORS JT. Hypothesized origin of microbial life in a prebiotic gel and the transition to a living biofilm and microbial mats. Comptes rendus biologies 2011;334:269-72.

[57] MARTIN W, et al. Hydrothermal vents and the origin of life. Nat Rev Micro 2008;6:805-14.

[58] TRAFICANTE DD. Optimum tip angle and relaxation delay for quantitative analysis. Concepts Magn Reson 1992;4:153-60.

CHAPTER 4:
ORIGIN OF GLYOXYLATE

In Chapters 2 and 3 I have discussed the role of glyoxylate in the formation of intermediates of the citric acid cycle, sugars, and sugar acids. While these results combine to indicate glyoxylate may have had an important role in prebiotic chemistry, no plausible prebiotic synthesis of glyoxylate has been identified. This may be due to the fact that detection of glyoxylate can be complicated by a tendency to disproportionate to glycolate and oxalate through a Cannizzarro reaction[1]. Both glycolate and oxalate however have been detected in the Murchison meteorite[2], suggesting glyoxylate may have been formed as well. Glycolic acid and oxalic acid are also both observed when insoluble HCN polymers, a byproduct of the HCN chemistry discussed in Chapter 1, are refluxed under acidic conditions[3], again suggesting that glyoxylate may be an intermediate product in this hydrolysis. In fact, the HCN dimer, imino-acetonitrile, and glyoxylate have the same oxidation state, making production of glyoxylate by hydrolysis of imino- acetonitrile a reasonable proposal. This pathway to glyoxylate is confounded by the tendency of imino-acetonitrile to dimerize to diaminomaleonitrile (DAMN) under hydrolysis conditions, but it is possible that glyoxylate could be produced by this mechanism. Alternately, diaminomaleonitrile has been shown to hydrolyze to DHF in the presence of titanium dioxide[4], which can undergo a cyanide promoted retro-aldol reaction to glyoxylate[5]. While titanium dioxide is not thought to be a prebiotically relevant mineral, if similar chemistry could be identified in the presence of another mineral species this pathway would be reasonable as a means of glyoxylate formation.

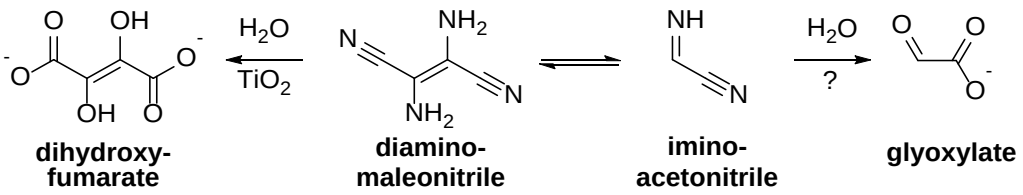


Figure 63 - Imino-Acetonitrile Chemistry Dimerization and potential hydrolysis of imino-acetonitrile (cyanide dimer).

This combination of tangential evidence for production of glyoxylate is sufficient grounding to reasonably consider the prebiotic role of glyoxylate; however, it is my belief that identifying a plausible synthesis of glyoxylate would help to cement its place in the prebiotic milieu. To this end, I have elected to investigate the reaction of cyanide, formaldehyde, and CO_2 to form the cyanohydrin of glyoxylate. Once formed, this cyanohydrin can equilibrate to form free glyoxylate and cyanide. At this point the fate of the glyoxylate will be determined based upon pH. In a moderate pH (7-10) system, glyoxylate may undergo a glyoxylose type reaction, leading to the formation of linear sugars, as discussed in Section 3.1. Alternately, under very high pH conditions, a gloxoin type reaction could occur, either leading to tartrate (Chapter 2), or in the presence of any free formaldehyde or aldoses, to sugar acids (Section 3.2).

4.1 Reaction of Glycolonitrile and CO_2

To investigate the reaction of cyanide, formaldehyde, and CO_2, initial reactions of glycolonitrile with CO_2 in the presence of sodium carbonate were conducted under 600 psi of CO_2. These conditions are reflective of conditions available on modern Earth at hydrothermal vents, where very high pressure plumes of CO_2[6] interact with seawater at basic conditions[7]. This is also likely representative of conditions across the entire planet during the initial formation of the oceans[8]. Glycolonitrile is simply the

cyanohydrin of formaldehyde, which forms readily when the two are combined under ambient aqueous conditions[9].

We conducted this reaction with the expectation that the glycolonitrile would react similarly to phenylacetonitrile. Phenylacetonitrile has been shown to react with cyanide in supercritical ethane in the presence of carbonate[10]. This reaction takes place by way of the reaction shown in Figure 64 resulting in the formation of a cyanoester product species. While this reaction is more complex than the intended reaction of glycolonitrile with carbon dioxide, it does illustrate the principle of a cyanocarbanion reacting with carbon dioxide to form a stable acid species.

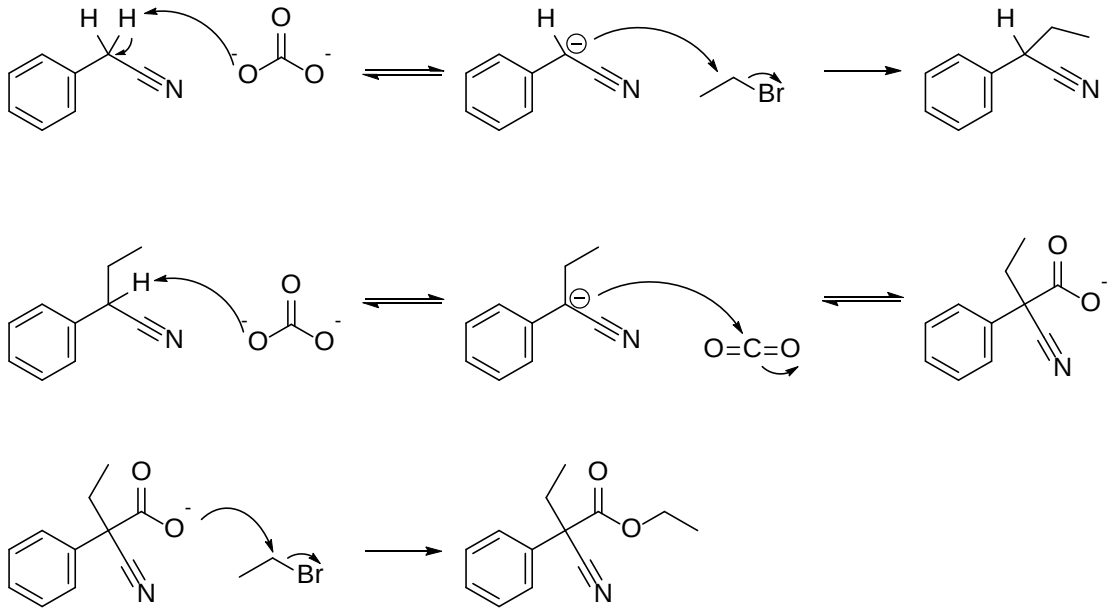


Figure 64 - Simplified Schematic of Cyanoester Formation from Phenylacetonitrile, Carbon Dioxide and 1 Bromo-Ethane Promoted by Carbonate
This mechanism takes place in a multiphase reaction mixture. For simplicity, and because our analogous reaction is expected to take place entirely in the aqueous phase, complexes with phase transfer catalysts have been omitted. The first step in this mechanism is the deprotonation of phenylacetonitrile by carbonate. This carbanion species then displaces bromide from 1-bromoethane in a simple SN2 reaction. This 1-phenyl-1-cyano propane product is then deprotonated by carbonate, with the resulting carbanion reacting with CO_2 to form 2-phenyl-2-cyano butanoate. This species displaces an additional bromide from 1-bromoethane in a second SN2 reaction.

These reactions were conducted by adding 10 mL of 2.0M sodium carbonate to a glass Parr liner and freezing for 24 hours. To this frozen carbonate solution, 10mL of 50wt% glycolonitrile solution was added. The glass liner was then inserted into the Parr reactor body, the reactor was sealed, flushed with CO_2 for 5 minutes, and pressurized to 400 psi. The heating mantle was then set to 40^{o}C, resulting in a final pressure of ~600 psi. The purpose of this experimental procedure was to limit the effect of base on the glycolonitrile until after the system had been flushed and pressurized. The reaction was then allowed to proceed for one week at 40^{o}C at which time it was vented and sampled for analysis by ^{13}C NMR. Upon removal from the Parr reactor, the aqueous phase had adopted a peach color, and contained numerous solid inclusions, most of which had the appearance of cloudy ice, however a few appeared as a black tar substance. These species respectively were likely a mixture of carbonate and potentially CO_2 clathrate, and glycolonitrile polymer, respectively. The substance with the appearance of ice was observed to decrease in volume over time.

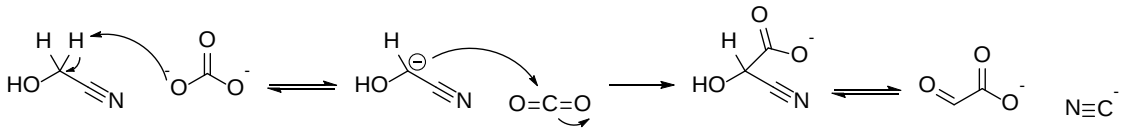


Figure 65 - Reaction of Glycolonitrile with CO_2 Under Basic Conditions.
Observation of ^{13}C-NMR peaks consistent with glyoxylate and glyoxylate cyanohydrin confirm that glycolonitrile can react analogously to phenylacetonitrile (**Figure 64**)

The ^{13}C NMR spectrum of this reaction was largely unreacted glycolonitrile (Figure 67), Increased magnification revealed peaks consistent with the hydrated cyclodimer of glycolonitrile (Figure 66), the known major product of the polymerization of glycolonitrile[11]. Some of the smaller peaks may be polymers of the cyclodimer. On close inspection peaks at 67, 125, and 176 ppm appear to be consistent with previously observed peaks for the cyanohydrin of glyoxylate, but are of very low intensity.

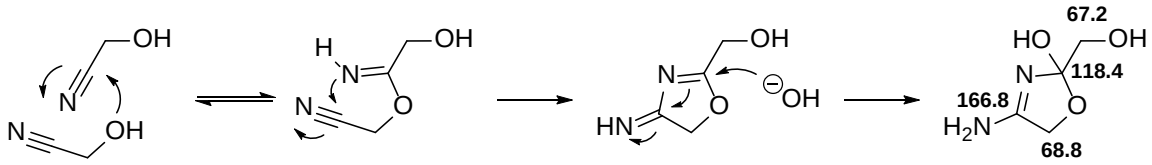


Figure 66- Mechanism of Cyclodimer Formation of Glycolonitrile.
The mechanism for the formation of the hydrated cyclodimer of glycolonitrile as proposed by Arrhenius et al. along with NMR assignments in d-DMSO for the product.

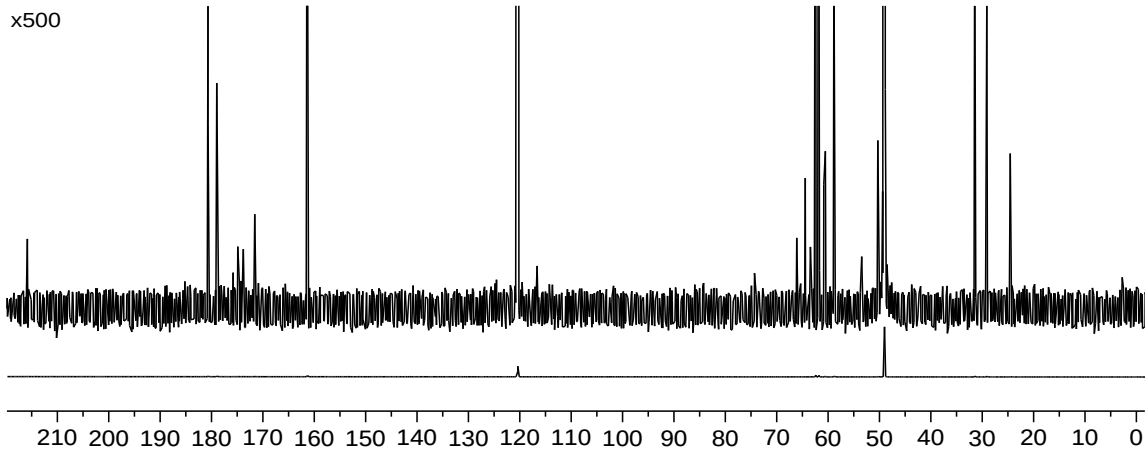


Figure 67 - ^{13}C NMR of the Reaction of 20mL 5.0M Glycolonitrile with CO_2 in the Presence of 20mmol of Na_2CO_3. After 10 days under 700 psi CO_2 and 40°C the two major peaks in the ^{13}C NMR spectrum are those of glycolonitrile at 120.4 and 49.0 ppm. Increased magnification shows that a number of species have formed. Very small peaks near 67, 125, and 176 ppm are consistent with previously observed peaks for the cyanohydrin of glyoxylate.

Having obtained an initial suggestion that reaction of glycolonitrile and CO_2 was possible, we sought to identify the conditions which would promote reaction to a larger extent than indicated by the very small ^{13}C NMR peaks observed. We chose first to investigate the effects of different counter ions on the reaction system. However, the design of the Parr reactor is such that only a single reaction can be conducted. With the length of these experiments and the potential variations in experimental conditions, this would make a direct comparison of reaction results time consuming and difficult. To overcome these limitations, we created carousel vial racks which fit inside of the Parr body, to allow 12 reactions to be conducted simultaneously. This “Parallel Parr Reactor” can hold two 6 vial carousels stacked atop each other, with a spacer between. A central shaft allows a mechanical impeller affixed with neodymium magnets to provide stirring to both levels. Reactions conducted in this reactor are thus held under identical temperature and pressure, at the expense of heat transfer.

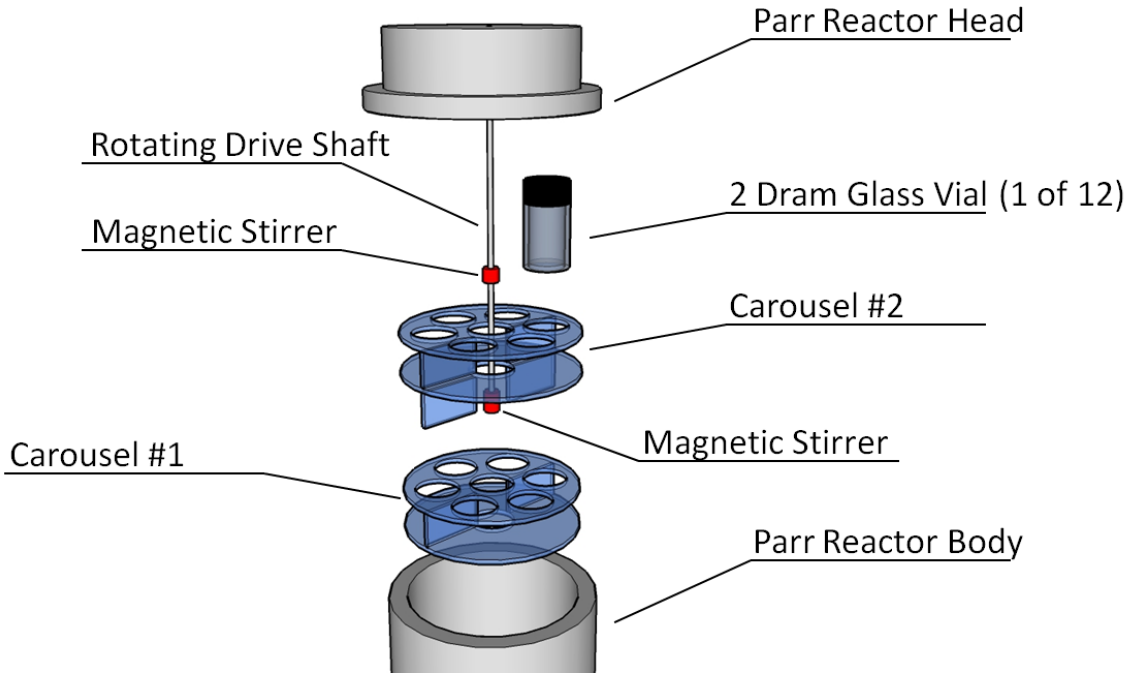


Figure 68 – Schematic of the Parallel Parr Reactor
To facilitate comparative studies of glyoxylate cyanohydrin formation, an existing Parr Reactor was modified to allow twelve reactions to be conducted simultaneously. Two carousel inserts were constructed of solvent welded Lexan, each capable of holding six vials, two drams in volume. The the stir blades on the mechanical drive shaft of the Parr were removed, and the blades were replaced with neodymium magnets to allow magnetic stirring of the vials. A second set of neodymium magnets was mounted on the shaft at the height of the second carousel to supply stirring to those vials.

Three sets of reactions have been carried out in the parallel Parr reactor. The first set of reactions tested the effect of different alkali metal cations along with calcium and ammonium as the counter ions to the carbonate base. Specifically, the counter ions tested were Li^+, Na^+, K^+, Cs^+, Ca^{2+}, and NH_4^+. The reactions were conducted by creating a stock solution of each carbonate (with the exception of Li_2CO_3 and $Ca(CO_3)_2$) and adding 1 mL of the stock solution to a 2 dram vial. For the insoluble lithium and calcium carbonates, 2 mmol were weighed into a 2 dram vial and 1 mL of water was added. For each carbonate two reactions were prepared. All vials were frozen for 24 hours. To each vial 1 mL of 50wt% glycolonitrile was added. The vials were lightly capped to prevent aerosol

contamination while allowing CO_2 to enter the reaction vessel. The vials were placed in the Parr carousels, and the reactor was assembled and purged for five minutes before being pressurized to 600 psi. These reactions were allowed to proceed for 8 days, at which time they were vented, and samples prepared for LCMS analysis immediately. The second set of reactions was prepared analogously, but with $MgCO_3$, $CaCO_3$, $SrCO_3$, $LiOH$, KOH, and NaOH as the bases. The third set of reactions consisted of a control where no base was added, and reactions of 1.0M Li_2CO_3 and 0.5M glycolonitrile, 1.0M $MgCO_3$ and 0.5M glycolonitrile, and finally a reaction of 1.0M CaCO3 with 5.0M glycolonitrile simply for comparison with the other two reaction sets.

Analysis was conducted using a Shimadzu LCMS and triple quad mass spectrometer. A multiple reaction monitoring (MRM) method was created and optimized using glyoxylate cyanohydrin prepared by combining glyoxylate with a 50% excess of cyanide in a pH 10.5 carbonate buffer to maximize similarity to reaction conditions. In MRM the mass spectrometer selects for a specific mass in the first quadrupole allowing all species of that mass to pass to the second quadrupole. These species are then fragmented by collision with nitrogen gas at a given energy, causing them to fragment. The spectrometer monitors for collision products in the third quadrupole. These fragmentation products serve as a molecular fingerprint, allowing both confirmation and quantitation of the parent species based on fragment intensity. A schematic of this method can be seen in Figure 69. Yields as determined by this method are presented in **Table 6**.

Table 6 – Yields of the Reaction of Glycolonitrile and Carbon Dioxide in the Presence of Carbonate
Reactions were conducted simultaneously with two replicates. Carbonate was present in 1M concentration and 25wt% glycolonitrile. Reactions were pressurized to 600 psi with CO_2 and allowed to react for 8 days. Yield was determined by MRM monitoring using a triple quadrupole LCMS with the final value being the average of three injections. Absent values were below the limit of detection.

	Time (days)	Base	Yield	Std. Dev.
1st Reaction Set	8	Cs_2CO3	0.001080%	0.000029%
	8	Cs_2CO3	0.001250%	0.000019%
	8	K_2CO_3	0.001770%	0.000014%
	8	K_2CO_3	0.000970%	0.000017%
	8	Na_2CO_3	0.002180%	0.000045%
	8	Na_2CO_3	0.002750%	0.000022%
	8	Li_2CO_3	0.011840%	0.000104%
	8	Li_2CO_3	0.010350%	0.000037%
	8	$CaCO_3$	0.000680%	0.000012%
	8	NH_4HCO_3	0.011500%	0.000045%
2nd Reaction Set	7	$MgCO_3$	0.000820%	0.000003%
	7	$MgCO_3$	0.000330%	0.000003%
	7	$CaCO_3$	0.000790%	0.000021%
	7	$CaCO_3$	0.000700%	0.000017%
	7	$SrCO_3$	0.000470%	0.000013%
	7	$SrCO_3$	0.000370%	0.000012%
	7	LiOH	0.006070%	0.000103%
	7	LiOH	0.010130%	0.000063%
	7	KOH	0.002410%	0.000011%
	7	KOH	0.002240%	0.000072%
	7	NaOH	0.003210%	0.000061%
	7	NaOH	0.003070%	0.000020%
3rd Reaction Set	8	Control (no base)	0.000130%	0.000010%
	8	Control (no base)	0.000130%	0.000017%
	8	$MgCO_3$ (low glycolontrile)	---	---
	8	$MgCO_3$ (low glycolontrile)	---	---
	8	Li_2CO_3 (low glycolonitrile)	0.008900%	0.000165%
	8	Li_2CO_3 (low glycolonitrile)	0.006700%	0.000095%
	8	$CaCO_3$	0.000650%	0.000023%
	8	$CaCO_3$	0.000500%	0.000009%

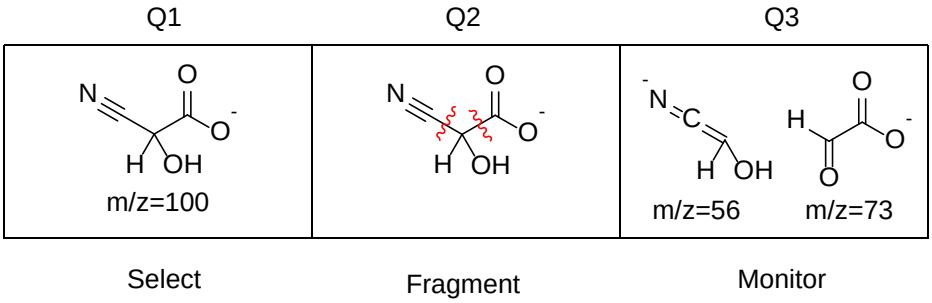


Figure 69 - Schematic Depiction of Multiple Reaction Monitoring Method for Determination of Glyoxylate Cyanohydrin. In this method, the glyoxylate cyanohydrin is allowed to pass through the first quadrupole of the mass spectrometer, excluding all other species. In the second quadrupole it is fragmented by collision with nitrogen gas, creating a "fingerprint" of fragmentation products specific to the cyanohydrin. In the third quadrupole these fragments are monitored, allowing quantitation of initial concentration.

Lithium carbonate and ammonium carbonate showed the highest yields of these reactions at .011%. While this is an incredibly small yield, it is the first plausible demonstration that glyoxylate could be produced prebiotically. This result is especially interesting when compared to the control reaction, which shows that the presence of lithium or ammonium carbonate causes a 100-fold increase in yield. While the likelihood of large concentrations of lithium having been present in the prebiotic ocean may not be high, cesium, sodium, and potassium carbonates also showed small amounts of glyoxylate cyanohydrin production, and ammonium carbonate shows a (relatively) large amount of glyoxylate cyanohydrin production, showing that lithium is not required. Magnesium, calcium, and strontium carbonates also show some catalytic activity, but not to the same degree demonstrated by the alkali metal cations.

Upon consideration of the yield data, what is immediately apparent is a trend of decreased yield with increased ionic radius/molecular weight in the alkali metals. This trend however breaks down when the alkaline earth metal cations are considered (Figure 70). In fact when all results all considered simultaneously, it appears as though lithium is an exceptional result with the other cations being roughly equal.

As smaller metal cations are stronger Lewis acids, we considered that perhaps, being the smallest cation, this was the factor which differentiated lithium from the other species. To test this theory, we used Pearson absolute hardness parameter[12] as a thermodynamic measurement of Lewis acidity. Pearson absolute hardness is equal to one half of the HOMO-LUMO gap in a species and consequently indicative of the ability to donate or accept electrons[13]. When the yield is plotted against the Pearson absolute hardness, a more robust linear relationship between yield and hardness is revealed for the group one cations, however the group two cations, specifically magnesium, break this trend (Figure 71).

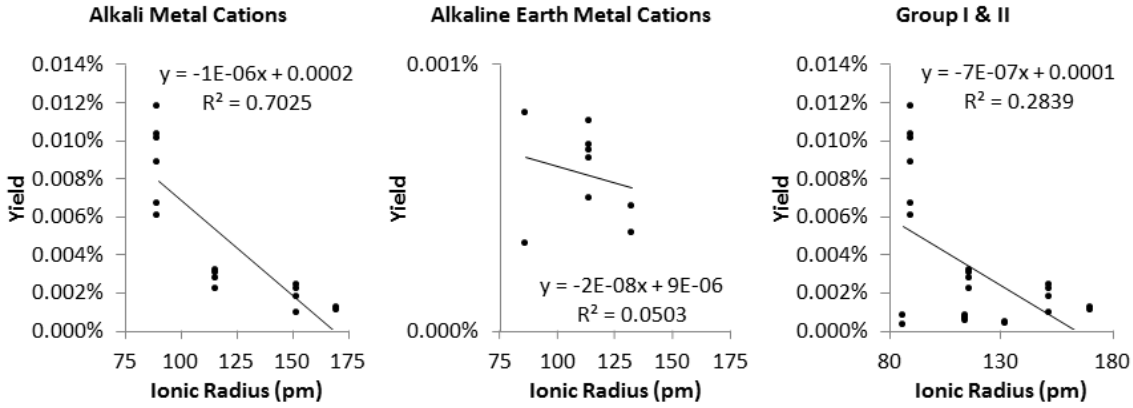


Figure 70 - Yield of the Reaction of Glycolonitrile with 600psi CO_2 and 2.0M Carbonate with Indicated Counter Cation plotted against Cation Ionic Radius Reaction yields of reactions with alkali metal cations are seen to increase exponentially with decreasing ionic radius, a trend which suggests the cation to be acting as a Lewis acid in the reaction mechanism. However when alkaline earth metals are considered, only a slight increase in yield with decreasing radius is observed. Taken together, the results appear as though lithium is somehow exceptional.

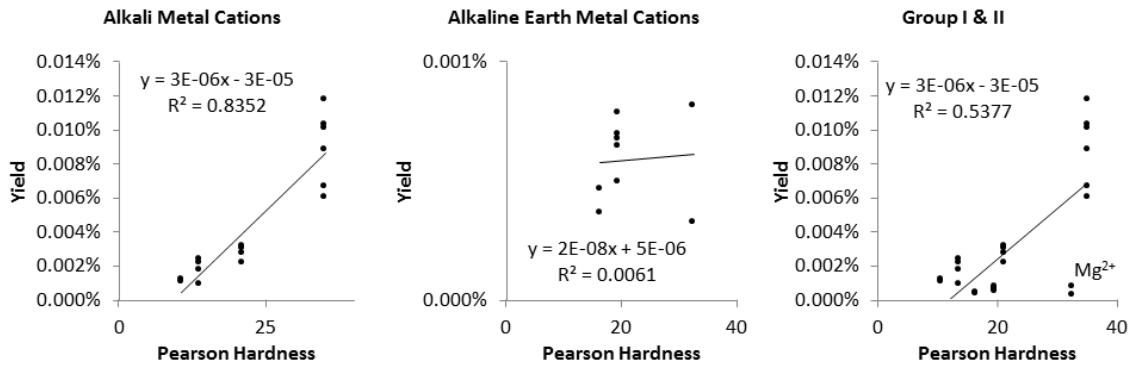


Figure 71 - Yield of the Reaction of Glycolonitrile with 600psi CO_2 and 2.0M Carbonate Plotted Against Cation Absolute Hardness. For reactions with alkali metal cations, yields are seen to increase as the absolute hardness of the cation increases. Magnesium breaks this trend for the group two metals.

Solubility of the different carbonates was considered as another potential differentiating factor, however no relationship was apparent using this parameter. In fact, all reactions were observed to have a considerable amount of solids upon removal from the parallel Parr reactor, suggesting that the formation of carbonate from CO_2 may have decreased the concentrations of the cations in the solution. This lead to the consideration that perhaps all cations were only available in the solution in very small amounts. If all species were available at a similar low concentration, and had a similar effect on the reaction, perhaps the differences are simply due to the mass of the cations. Collision theory dictates that for two reactions that differ only by molecular weight of a reactant, the rate of the reaction (r_{AB}) should scale with the reduced mass of the collision pair as:

$$r_{AB} \sim \sqrt{\frac{m_A \cdot m_B}{m_A + m_B}}$$

where m_A and m_B are molecular weights of the reactants. When this mass scaling factor of r_{AB} is plotted against yield, the best relationship thus far between yield and cation properties is obtained (, however some discrepancies remain.

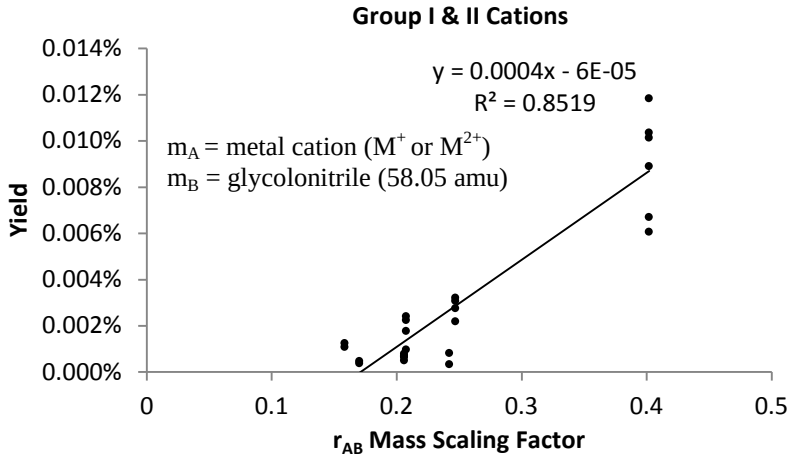


Figure 72 – Relationship between Cation Reduced Mass and Yield

In aqueous solution calcium carbonate is known to exist as an equilibrium between solid $CaCO_3(s)$, dissociated Ca^{2+} and CO_3^{2-}, and a solvated ion pair $CaCO_3^{+-}$[14]. Under high partial pressures of CO_2, this equilibrium would be driven to $CaCO_3(s)$ and $CaCO_3^{+-}$. If this same behavior is assumed for the other group two carbonates, and the reduced mass of the metal carbonate solvated ion pair is used for the same type of reduced mass calculation the strongest relationship this far is revealed. This relationship suggests that these metal cations increase the rate of reaction simply through charge stabilization, rather than a unique electronic property. This charge stabilization could either be through the formation of an ion pair with glycolonitrile, or CO_2 as shown in Figure 73 Unfortunately, the mass of glycolonitrile and CO_2 are too close together, and no insight can be garnered into the mechanism of action (Figure 74).

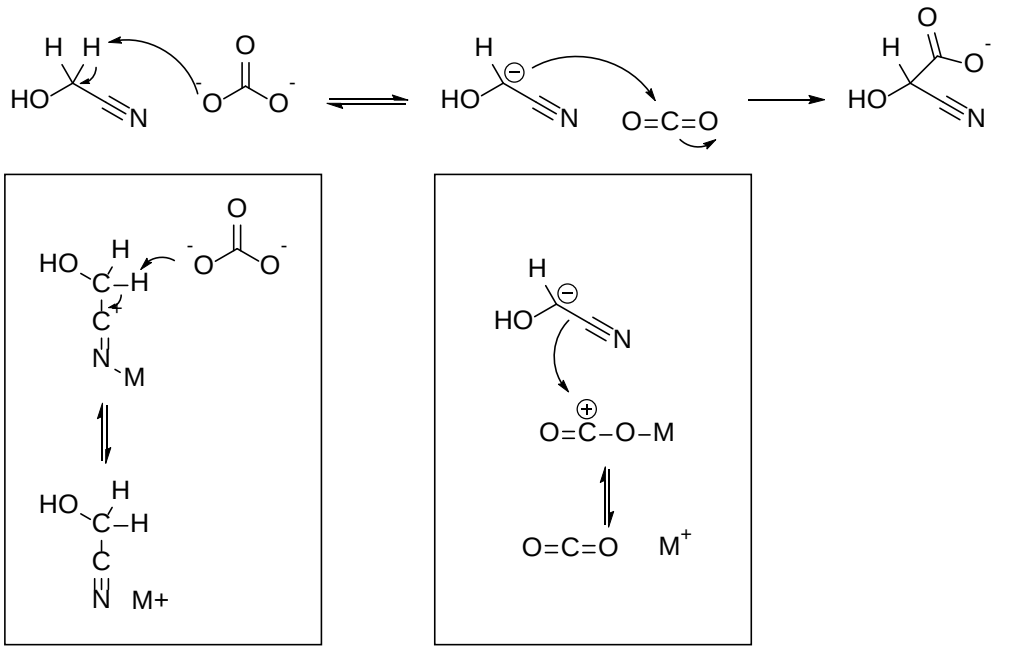


Figure 73 – Two Potential Catalytic Complexes Formed with Metal Cations. Positively charged metal cations could either facilitate in the deprotonation of glycolonitrile, or in polarizing the CO_2 bond to facilitate attack by glycolonitrile.

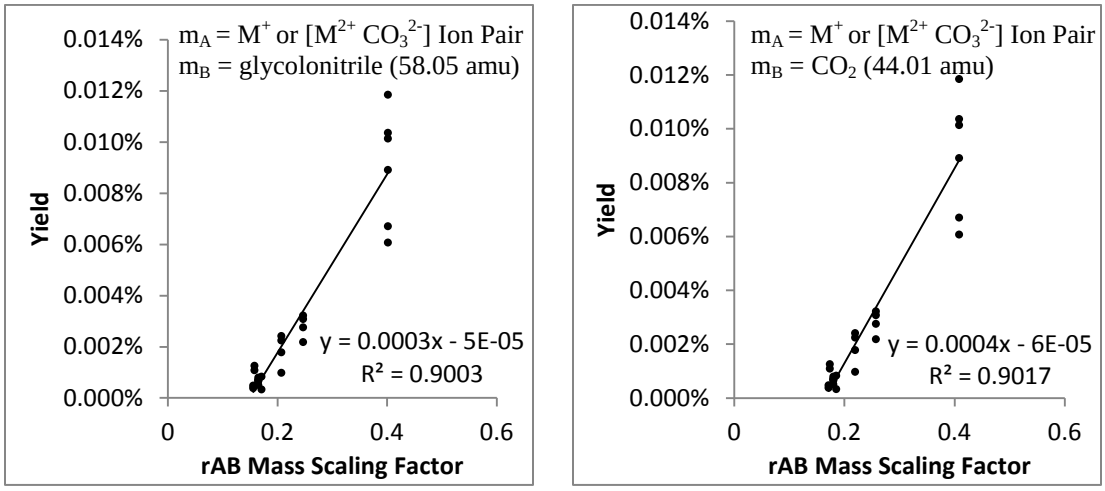


Figure 74 – Relationship between M^+ or [M^{2+} CO_3^{2-}] Ion Pair Reduced Mass and Yield

4.2 Summary

We have demonstrated that under conditions of room temperature and 600 psi CO_2 that reaction of glycolonitrile with carbon dioxide to form the cyanohydrin of glyoxylate at very low yields. This rate of this reaction is increased 100-fold by the presence of lithium carbonate, and increased to a lesser degree by other metal cations. Comparison of the reaction yields suggests that charge stabilization may be the mechanism of catalysis. If this is the case, utilization of salts which will remain soluble in the presence of high partial pressures of CO_2 may be productive to increase reaction yields.

Of all the species tested, ammonium carbonate is likely the most attractive catalyst from a prebiotic perspective. Ammonia is a species expected to be abundant on early Earth [15-17], and its catalytic activity was equal to that of lithium carbonate. Unlike metal carbonates however, ammonia/ammonium can undergo chemical as well as physical interactions requiring a greater level of monitoring, difficult to provide at the low yields obtained.

While the yields presented herein were quite low, the conditions utilized were quite mild within the scope of possible hydrothermal conditions, allowing opportunity for improvement. Simply increasing temperature, may be enough to increase yield, and given that hydrothermal systems have been found with temperatures up to 333^{o}C, there is ample room to explore realistic conditions. Alternately, super critical CO_2 ($scCO_2$) may be a system worth considering. While this may seem exotic, conditions in hydrothermal faults

may be sufficient to produce $scCO_2$ and chemistry in such an environment has been proposed to be a contributor to prebiotic chemistry[18]. As the reaction of phenylacetonitrile with CO_2 that inspired this work was conducted in $scCO_2$ this may also be a viable direction for continuation of this work.

4.3 Experimental

LCMS (MRM)- A Shimadzu LC-MS 8030 was used for quantitative detection of the glyoxylate cyanohydrin. Samples were diluted 1000:1 in LCMS H_2O and injected without a column, and multiple reaction monitoring (MRM) was used for quantitation. Using the following settings.

Precursor (m/z)	Product (m/z)	Pause time (msec)	Dwell time (msec)	Q1 Pre-bias (V)	CE	Q3 Pre-bias (V)
100	73.2	3.0	10.0	21	10	15
100	56.3	3.0	10.0	21	15	11
100	59.35	3.0	10.0	21	10	12

Reactions of Glycolonitrile and Carbonate Species: In individual 2-dram vials, 1.0 mL of LCMS H_2O was combined with 2 mmol of solid carbonate, labeled and frozen for 24 hours. This was done to prevent interaction of the glycolonitrile with the base until after dissolved CO_2 was present. 1 mL of 50 wt% glycolonitrile was added to each reaction mixture atop the frozen base solution. The vials were lightly capped and placed into the parallel Parr carousels. This was to prevent aerosol contamination while allowing CO_2 to enter the vial. The Parr reactor was then sealed and purged with carbon dioxide for five minutes. After that time, the vessel was pressurized up to 600 psi of carbon dioxide atmosphere and was monitored with the constant pressure capabilities of an ISCO 500 mL tank capacity high pressure syringe pump.

4.4 References

[1] CANNIZZARO S. Ueber den der Benzoësäure entsprechenden Alkohol. Justus Liebigs Annalen der Chemie 1853;88:129-30.

[2] PELTZER ET, et al. The chemical conditions on the parent body of the murchison meteorite: Some conclusions based on amino, hydroxy and dicarboxylic acids. Advances in Space Research 1984;4:69-74.

[3] RUIZ-BERMEJO M, et al. New Insights into the Characterization of 'Insoluble Black HCN Polymers'. Chem Biodivers 2012;9:25-40.

[4] SALADINO R, et al. Photochemical synthesis of citric acid cycle intermediates based on titanium dioxide. Astrobiology 2011;11:815-24.

[5] BUTCH C, et al. Production of Tartrates by Cyanide-Mediated Dimerization of Glyoxylate: A Potential Abiotic Pathway to the Citric Acid Cycle. J Am Chem Soc 2013;135:13440-5.

[6] LUPTON J, et al. Submarine venting of liquid carbon dioxide on a Mariana Arc volcano. Geochemistry, Geophysics, Geosystems 2006;7:Q08007.

[7] MOTTL MJ. Highest pH. Geochem News 2009;141.

[8] PINTI DL. The origin and evolution of the oceans. Lectures in Astrobiology: Springer; 2005. p. 83-112.

[9] SLEEP NH, et al. H2-rich fluids from serpentinization: Geochemical and biotic implications. Proceedings of the National Academy of Sciences of the United States of America 2004;101:12818-23.

[10] WHEELER C, et al. Phase-Transfer-Catalyzed Alkylation of Phenylacetonitrile in Supercritical Ethane. Industrial & Engineering Chemistry Research 2002;41:1763-7.

[11] ARRHENIUS G, et al. Glycolonitrile Oligomerization: Structure of Isolated Oxazolines, Potential Heterocycles on the Early Earth. The Journal of Organic Chemistry 1997;62:5522-5.

[12] PEARSON RG. Absolute electronegativity and hardness: application to inorganic chemistry. Inorganic Chemistry 1988;27:734-40.

[13] ARNETT EM, LUDWIG RT. On the relevance of the Parr-Pearson principle of absolute hardness to organic chemistry. J Am Chem Soc 1995;117:6627-8.

[14] LAFON GM. Calcium compleing with carbonate ion in aqueous solutions at 25°C and 1 atmosphere. Geochimica et Cosmochimica Acta 1970;34:935-40.

[15] SUMMERS DP, CHANG S. Prebiotic ammonia from reduction of nitrite by iron (II) on the early Earth. 1993.

[16] BRANDES JA, et al. Abiotic nitrogen reduction on the early Earth. Nature 1998;395:365-7.

[17] HO PT, TOWNES CH. Interstellar ammonia. Annual Review of Astronomy and Astrophysics 1983;21:239-70.

[18] SCHREIBER U, et al. Hypothesis: Origin of Life in Deep-Reaching Tectonic Faults. Origins of Life and Evolution of Biospheres 2012;42:47-54.

CHAPTER 5:
CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

The work presented in this document presents a detailed view of reactions of glyoxylate which could potentially contribute to the formation of biologically relevant molecules on an abiotic earth. High pH reactions of glyoxylate show potential as an originating point for many important metabolic intermediates. The simple and efficient reaction of glyoxylate and cyanide to produce tartrate is robust and occurs under a variety of conditions. While the required high hydroxide concentrations are considered to be extreme (and unusual) in conventional prebiotic line of thought, there are at least three plausible prebiotic scenarios yielding high pH which one can consider. The first possible scenario is the widely investigated alkali hydrothermal vent system[1-3] which in modern systems have been measured up to pH 12.7[4]; however, some type of containment, such as smoker chimneys[5], would be necessary to help mitigate the drawbacks of hydrolysis and dilution endemic to this system[6]. Related, and perhaps even more attractive, reservoirs of natural hydroxide are lakes fed by alkali springs,[7] which could act as a natural reactor for this type of chemistry. Such a lake would have the advantage of containment, and lower temperature mitigating the problems faced by oceanic vents, while simultaneously allowing access to photolytic production of glyoxylate[8]. Another possible high alkaline environment can be found in the interlayer framework of double-layered hydroxides (e.g. hydrotalcites), which are especially conducive for uptake and concentration of anions (such as glyoxylate and cyanide) from dilute solutions[9]. The combination of these possible environments offers sufficient potential for highly alkaline

environments on early earth that need to be experimentally validated from a prebiotic perspective, for the chemistry described here.

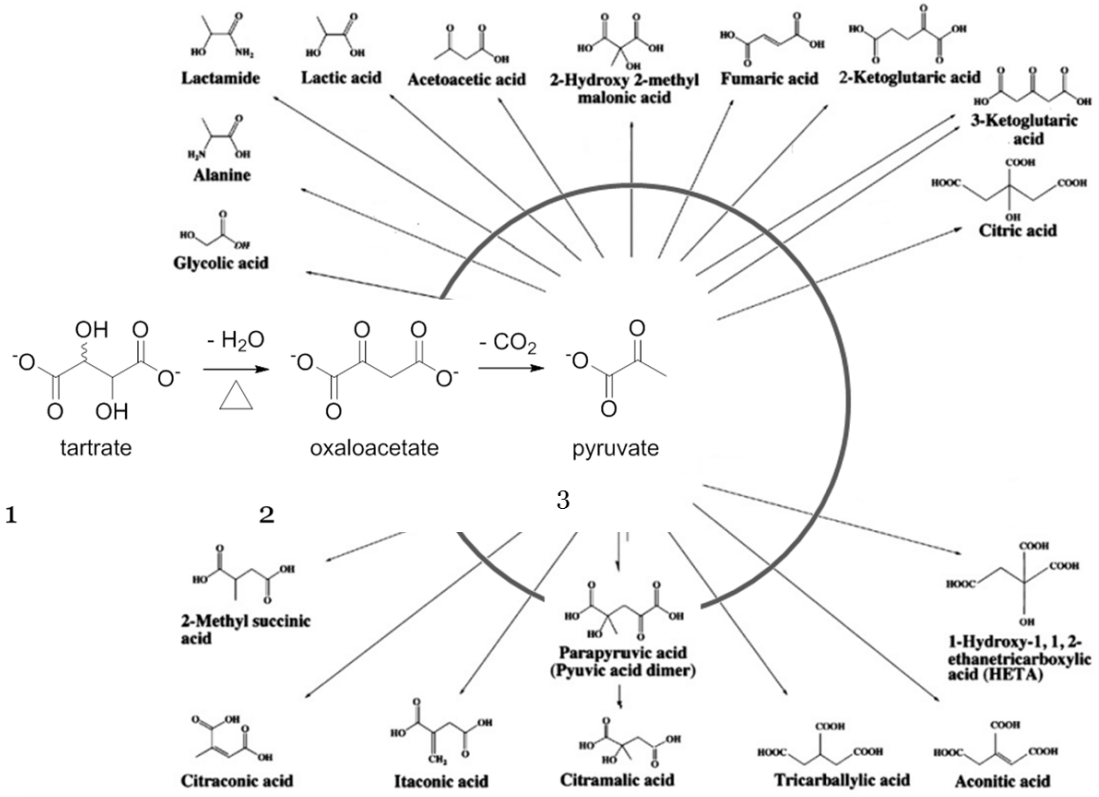


Figure 75 – Chemical Pathway Leading from Tartrate to Many Intermediates of the Citric Acid Cycle. 1) The formation of oxaloacetate, an intermediate of the citric acid cycle, from the known thermal dehydration of tartrate is a simple and expeditious path from the chemistry of glyoxylate into an early metabolic scenario. 2) Oxaloacetate is known to decarboxylate to pyruvate, 3) reactions of which have been shown to lead to most of the intermediates of the citric acid cycle. This type of chemistry could potentially lay the groundwork for a system similar to the citric acid cycle to arise from these components.

Assuming high pH conditions were available on the early earth, the ability of the glyoxoin reaction to proceed even at low temperature, low concentration, and low solubility suggests that with a reliable source of glyoxylate the reaction is quite likely to occur. This reliability, combined with known chemistry leading from tartrate to oxaloacetate[10, 11] to pyruvate[12], and from pyruvate to many of the intermediates of the citric acid cycle[13] illustrates how glyoxylate and cyanide could be a stepping off point to lead to a rudimentary metabolic system (Figure 75). A reductive citric acid cycle

at hydrothermal vents has been proposed as a potential early metabolism [14-16]; chemistry of the glyoxylate scenario[17] under these conditions may be sufficient to create an environment capable of launching such a chemical system.

In addition to access to the citric acid cycle as described above, the high pH chemistry of dihydroxyfumarate, and thus glyoxylate, also gives access to aldonic acids and glycerate by reaction with formaldehyde, glycolaldehyde, and aldoses. Production of aldonic acids from aldehydes under high pH conditions is a potential means of bypassing a major flaw in the formose reaction, the instability of the sugar products[18]. The formose reaction is known to produce sugars at high pHs where the sugar products are unstable[19]. Thus it is possible that the formose reaction could take place in an environment where reaction of DHF and aldoses to form aldonic acids could also occur. This could create a scenario where selective reaction of DHF with aldose products of the formaldehyde reaction removing them from the reaction pathway, and protecting them against decomposition.

The aldonic acid products of the reaction of DHF and aldoses can act as more stable sinks for the storage of carbohydrates. These storage forms would then have to be reduced, or undergo oxidative decarboxylation to return to the sugar form, processes which merit further research. Alternately, these aldonic acid products could contribute interesting chemistry to the prebiotic landscape in their own rights. Conversion of glycerate to phosphoglycerate has been demonstrated by reaction with trimetaphosphate, and phosphoglycerate has been proposed as a potential starting point (originating from pyruvic acid from a reductive citric acid cycle) for gluconeogenesis[1]. In addition to this tie in with the potential for a prebiotic metabolism, higher aldonic acids could be

structural materials in their own rights. Oxidation of the terminal alcohol of hexonic acid yields hex-6-ulosonic acid, the cyclic form of which possesses a passing similarity to phosphoribose. If this species is capable of forming 2 - 4 hemiacetal linkages it could act as a potential backbone polymer for an informational system.

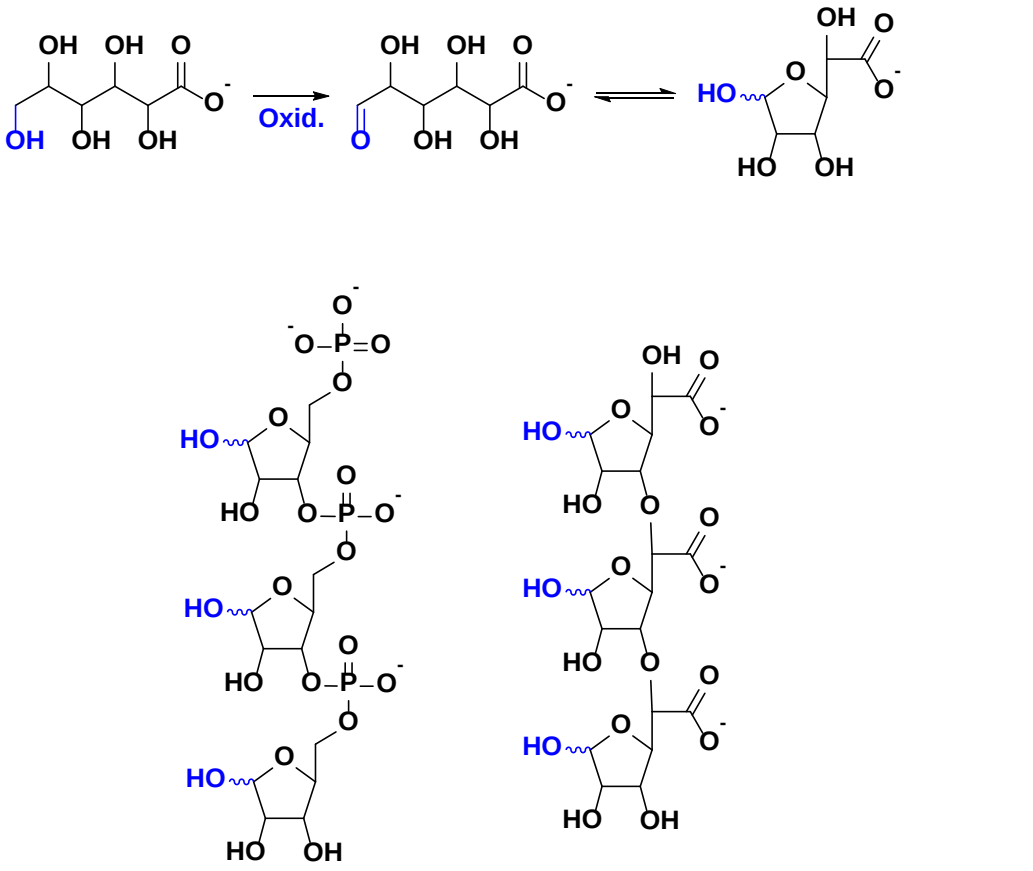


Figure 76 – Potential Conversion of Hexonic Acid to a Structure Capable of Acting as a Backbone for an Informational Polymer and Comparison to Polyphosphoribose. Oxidation of the terminal hydroxymethyl group of hexonic acid to hex-6-ulosonic acid creates a species capable of ring closure, and possessed of a charged sidegroup. If this species is capable of 2-4 linkages, it would bear a strong similarity to a ribose backbone.

Even should such high pH conditions be unavailable on early earth, the glyoxylose reaction as an alternative to the formose reaction gives demonstration to how glyoxylate may be an important source molecule for prebiotic reactions. Compared to the formose reaction the glyoxylose reaction has the advantages of linear product formation

and progressing at more neutral pH. This work dovetails nicely with the chemistry of the glyoxylate scenario established by Sagi et al. [20] which shows that reactions of dihydroxyfumarate with aldehydes can also produce sugars under these conditions. Again, these chemistries may occur side by side, and perhaps also alongside the formose reaction, creating a robust system for sugar formation, a necessary prerequisite to the arrival at the RNA world[21].

Together, the chemistries described above illustrate the potential impactful role glyoxylate could have played in the chemistry of Hadean Earth leading up to the advent of the RNA World. Taken together they are strong evidence that glyoxylate should be put under closer scrutiny by the prebiotic community, specifically with an eye toward potential origins of glyoxylate. Our work on the origin of glyoxylate provides a proof of concept, that glyoxylate can be created from cyanide, formaldehyde, and carbon dioxide. With continued study, better conditions for this reaction to take place may be discovered. Alternately with increased scrutiny by the prebiotic community, other methods may be identified for the formation of glyoxylate which would confirm these chemistries as important prebiotic mechanisms.

5.2 Recommendations

5.2.1 Glyoxoin Reaction

The major drawback of the glyoxoin reaction, as described, is the high pH required to activate the umpolung catalysis by cyanide. As such, efforts should be made to find conditions which abate this requirement. Double layered hydroxides (DLHs) were mentioned above, and provide an interesting environment in which the glyoxoin reaction could occur at high pH. DLHs are found in nature as hydrotalcite and similar

minerals[22] and can be formed readily in solution[23, 24] making them ideal candidates for further evaluation, from both a prebiotic and laboratory perspective. I recommend that reactions of glyoxylate and cyanide under near neutral bulk solvent conditions in the presence of basic DLHs be investigated to determine whether this high pH requirement may be bypassed.

Another means to potentially bypass the high pH requirement for the dimerization of glyoxylate is to use an umpolung catalyst other than cyanide. The more electron withdrawing the umpolung catalyst employed in this reaction, the lower the pH requirement should become. Two umpolung catalysts that have been demonstrated for benzoin condensations and are prebiotically plausible, are N-heterocyclic carbenes (NHCs) [25, 26], and phosphonates[27]. Phosphite as a potential catalyst for this reaction is particularly interesting although it may not be electron withdrawing enough, there is ample evidence that free phosphite was a significant component of the prebiotic ocean[28]. Investigating the catalytic potentional of phosphite and NHCs for the dimerization of glyoxylate in only mildly basic media may illuminate a means of entry into the lower pH chemistry of the glyoxylate scenario as described by Sagi et al.

The tartrate products of this reaction have not been thoroughly considered in the prebiotic literature. The known conversion of tartrate to oxaloacetate provides access to the citric acid cycle as described, and we have seen some evidence that allowing our reaction solutions to evaporate under mild heating is sufficient to drive this conversion. Given the compatibility between the tartrate forming reaction and hydrothermal vent conditions, invoking an evaporative dehydration may not be reasonable. An aqueous reaction capable of dehydrating tartrate would then be valuable to pushing into the citric

acid cycle. One such method may be by phosphorylation of tartrate. The phosphorylation of glycerate demonstrated by Orgel et al. could feasibly occur utilizing tartrate in the place of glycerate. The phosphate group would be a better leaving group than the hydroxide it replaced, and thus, deprotonation at the β position may eliminate phosphate, resulting in the formation of oxaloacetate after tautomerization. Such a transformation would demonstrate entry into the citric acid cycle in a manner more compatible with tartrate formation.

5.2.2 High pH Reaction of DHF and Aldehydes

Ascorbic acid (Vitamin C) contains a carboxy-ene-diol moiety analogous to that of DHF. Conducting reactions of ascorbic acid with formaldehyde and glycolaldehyde would be an interesting proof of concept to showing that the deoxalation pathway could occur using a molecule other than DHF. If the reaction can be generalized, further extension to a species such as isoesculetin[29], could provide a novel way of synthesizing phenyl carbohydrates.

Also interesting would be identifying conditions under which this reaction could be run in organic solvents. Our kinetic and computational investigations indicate that the deoxalation is likely driven by nucleohpilicity of hydroxide. While high hydroxide concentrations will not be obtainable in organic solvents, introduction of another small nucleophile such as fluoride, from a high concentration of CsF could be adequate to drive a fragmentation to product and oxylfluoride.

5.2.3 Glyoxylose Reaction

First among the unanswered questions around the glyoxylose reaction is whether it is truly autocatalytic in the manner of the formose reaction. Chemically, all the pieces are

present for an autocatalytic cycle based off of amplification by retroaldol reactions of higher sugars. In practice, the reaction of higher sugars with glyoxylate may lead to a stable long chain ulosonic acid which could prevent the retroaldol reaction. The propensity for retroaldol reactions to occur in this system could be tested by beginning a reaction with isotopically labeled fructose. We have shown this sugar will react with glyoxylate. By beginning with the isotopically enriched sugar, the reaction mixture can be monitored for retroaldol products such as glycolaldehyde and DHA, and by their enrichement, say with certainty fructose was the parent molecule.

Efforts should also be made to quantify and understand the kinetics of the glyoxylose reaction such that the effects of additives on sugars yields can be determined. Benner recently demonstrated the addition of borate to the formose reaction yielding an increased concentration of ribulose[30-32]. If a similar mineral effect occurs in the glyoxylose reaction, the preexisting tendency to form pentoses could be amplified.

5.2.4 Formation of Glyoxylate

Demonstrating a more robust formation of glyoxylate is necessary to remove any doubts about the significance of the glyoxylate scenario with regards to the chemistry of the origin of life. The reaction of glycolonitrile with CO_2 may accomplish this under more energetic conditions. I recommend investigation of reactions of lithium carbonate and ammonium carbonate both in aqueous solution at high temperature and under CO_2 pressure, as well as in super critical CO_2.

Should formation of the glyoxylate cyanohydrin from glycolonitrile prove inefficient, another simple reaction plausible under hydrothermal conditions exists, but is potentially quite difficult to conduct in the laboratory. Serpentinization is the process of

oxidation of olivine to serpentine and magnetite. This process is known in nature to reduce CO_2 to formate, formaldehyde, methanol and methane, and is expected to have been occurring widely on prebiotic earth and leads to the alkaline conditions at hydrothermal vents, invoked previously[33]. Substituting carbon dioxide with oxalic acid has the potential to yield glyoxylate as the reduction product (Figure 77). As oxalic acid has been observed numerous times under plausibly prebiotic conditions [34-36], it serves as a reasonable source molecule for glyoxylate. Since, serpentinization occurs readily only under conditions of very high temperature and pressure, finding a laboratory method to conduct or simulate this process could be a very productive avenue of exploration for investigating the formation of glyoxylate. The most straight forward method would be to heat an aqueous solution of olivine and oxalate to the limit of safety in a Parr or other high pressure reactor, and measure the output. As serpentinization requires temperatures in excess of 250^{o}C and aqueous environment investigation could rapidly challenge the pressure limits of conventional bomb reactors.

Figure 77 – Simple Scheme of the Serpentinization Reaction of Oxalate

If reaction of olivine and aqueous oxalate at temperature and pressure does not result in the reduction of oxalate, another option to simulate using an iron rich olivine electrode. Using an olivine electrode in an aqueous solution containing oxalate it may be possible to use cyclic voltammetry to simulate the serpentinization reaction may be a reasonable alternative.

5.3 References

[1] MARTIN W, RUSSELL MJ. On the origin of biochemistry at an alkaline hydrothermal vent. Phil Trans R Soc B 2007;362:1887-926.

[2] MARTIN W, et al. Hydrothermal vents and the origin of life. Nature Rev Microbiol 2008;6:805-14.

[3] GLANSDORFF N, et al. The origin of life and the last universal common ancestor: do we need a change of perspective? Res Microbiol 2009;160:522-8.

[4] MOTTL MJ. Highest pH. Geochem News 2009;141.

[5] HAYMON RM. Growth history of hydrothermal black smoker chimneys. 1983.

[6] MILLER SL, BADA JL. Submarine hot springs and the origin of life. Nature 1988;334:609-11.

[7] PEDERSEN K, et al. Distribution, diversity and activity of microorganisms in the hyper-alkaline spring waters of Maqarin in Jordan. Extremophiles 2004;8:151-64.

[8] MENOR-SALVÁN C, MARÍN-YASELI MR. A New Route for the Prebiotic Synthesis of Nucleobases and Hydantoins in Water/Ice Solutions Involving the Photochemistry of Acetylene. Chemistry – A European Journal 2013;19:6488-97.

[9] BOCLAIR J, et al. Cyanide Self-Addition, Controlled Adsorption, and Other Processes at Layered Double Hydroxides. Orig Life Evol Biosph 2001;31:53-69.

[10] CHATTAWAY FD, RAY FE. II.—The decomposition of tartaric acid by heat. Journal of the Chemical Society, Transactions 1921;119:34.

[11] DANIELÁCHATTAWAY F, DAVIDÁPARKES G. LXXIX.—The formation of derivatives of oxalacetic acid from tartaric acid. Journal of the Chemical Society, Transactions 1923;123:663-9.

[12] WILEY RH, KIM KS. Bimolecular decarboxylative self-condensation of oxaloacetic acid to citrolyformic acid and its conversion by oxidative decarboxylation to citric acid. The Journal of Organic Chemistry 1973;38:3582-5.

[13] COOPER G, et al. Detection and formation scenario of citric acid, pyruvic acid, and other possible metabolism precursors in carbonaceous meteorites. Proceedings of the National Academy of Sciences 2011;108:14015-20.

[14] WACHTERSHAUSER G. Origin of life in an iron-sulfur world. The molecular origins of life 1998:206-18.

[15] ANET FA. The place of metabolism in the origin of life. Current opinion in chemical biology 2004;8:654-9.

[16] WÄCHTERSHÄUSER G. Life as we don't know it. Science 2000;289:1307-8.

[17] ESCHENMOSER A. The search for the chemistry of life's origin. Tetrahedron 2007;63:12821-43.

[18] KOPETZKI D, ANTONIETTI M. Hydrothermal formose reaction. New Journal of Chemistry 2011;35:1787-94.

[19] WEISS AH, JOHN T. Homogeneously catalyzed formaldehyde condensation to carbohydrates: III. Concentration instabilities, nature of the catalyst, and mechanisms. Journal of Catalysis 1974;32:216-29.

[20] SAGI VN, et al. Exploratory Experiments on the Chemistry of the "Glyoxylate Scenario": Formation of Ketosugars from Dihydroxyfumarate. J Am Chem Soc 2012;134:3577-89.

[21] DWORKIN JP, et al. The roads to and from the RNA world. Journal of Theoretical Biology 2003;222:127-34.

[22] RIVES V. Layered double hydroxides: present and future: Nova Publishers; 2001.

[23] PITSCH S, et al. Concentration of simple aldehydes by sulfite-containing double-layer hydroxide minerals: Implications for biopoesis. Helvetica chimica acta 2000;83:2398-411.

[24] KRISHNAMURTHY R, et al. Formation of glycolaldehyde phosphate from glycolaldehyde in aqueous solution. Orig Life Evol Biosph 1999;29:333-54.

[25] SHIMAKAWA Y, et al. Facile route to benzils from aldehydes via NHC-catalyzed benzoin condensation under metal-free conditions. Tetrahedron Letters 2010;51:1786-9.

[26] BOULANGER E, et al. Photochemical Steps in the Prebiotic Synthesis of Purine Precursors from HCN. Angewandte Chemie International Edition 2013;52:8000-3.

[27] GLIGA A, et al. New Umpolung Catalysts: Reactivity and Selectivity of Terpenol-Based Lithium Phosphonates in Enantioselective Benzoin-Type Couplings. European Journal of Organic Chemistry 2011;2011:256-63.

[28] PASEK MA. Rethinking early Earth phosphorus geochemistry. Proceedings of the National Academy of Sciences 2008;105:853-8.

[29] BRANDT RB, et al. Potential transition-state inhibitors of glyoxalase. I. International Journal of Quantum Chemistry 1982;22:335-43.

[30] BENNER SA, et al. Asphalt, water, and the prebiotic synthesis of ribose, ribonucleosides, and RNA. Accounts of chemical research 2012;45:2025-34.

[31] KIM H-J, et al. Synthesis of carbohydrates in mineral-guided prebiotic cycles. J Am Chem Soc 2011;133:9457-68.

[32] RICARDO A, et al. Borate Minerals Stabilize Ribose. Science 2004;303:196.

[33] SLEEP NH, et al. H2-rich fluids from serpentinization: Geochemical and biotic implications. Proceedings of the National Academy of Sciences of the United States of America 2004;101:12818-23.

[34] SCHWARTZ AW, et al. Recent progress in the prebiotic chemistry of HCN. Origins Life Evol Biosphere 1984;14:91-8.

[35] MILLER S, et al. Origin of organic compounds on the primitive earth and in meteorites. Journal of Molecular Evolution 1976;9:59-72.

[36] MILLER SL. Current status of the prebiotic synthesis of small molecules. Chem Scr B 1986;26:5-11.

APPENDIX A
FORMATION OF GLYCINE POLYPEPTIDES FROM GLYOXYLATE BY TRANSAMINATION AND COUPLING PROMOTED BY HEXAMETHYLENETETRAMINE

Hexamethylenetetramine (HMTA), is a potential thermodynamic sink for formaldehyde and ammonia due to its ready formation ($K_{eq} > 10^{10}$) from the condensation of four equivalents of formaldehyde and six equivalents of ammonia[1]. Our initial investigations into HMTA were in search of a source of dimethylamine, motivated by a desire to explore the possibilities of the ionic liquid dimethylammonium dimethylcarbamate as an alternative prebiotic solvent. In these experiments we heated HMTA in the presence of 10M ammonium formate resulting in the decomposition of HMTA and formation of methylamine, dimethylamine, trimethyl amine, formamide, N-methylformamide, and N,N-dimethylformamide, as well as a very large signal corresponding to HMTA plus a methyl unit. We have proposed that this methyl-HMTA is formed by capture of an additional equivalent of formaldehyde (Figure 78).

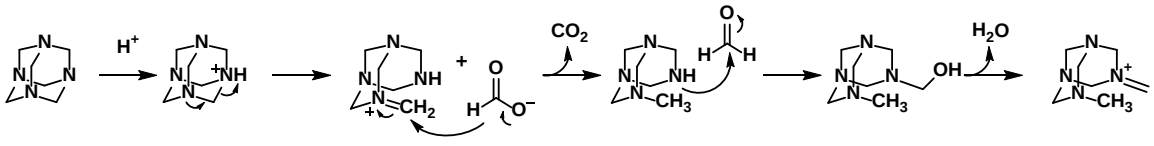


Figure 78 – Proposed Mechanism of Methyl-HMTA Formation. Formation of methyl-HMTA is initiated by partial HMTA ring opening, and reduction by formate. The newly formed secondary amine reacts with free formaldehyde formed by the reverse equilibrium of HMTA. This formyl-HMTA then undergoes a dehydration to yield methyl-HMTA.

Based on the affinity of HMTA for free formaldehyde, as well as the evidence that a redox process was taking place in the system, we postulated that the same process could occur with free glyoxylic acid, and potentially yield glycine. To this end I investigated the products of heating HMTA in a buffer of glyoxylic acid and found that glyoxylic acid was produced. I further investigated reactions of HMTA and pyruvic acid,

oxalacetic acid, and α-ketoglutaric acid, and in each case found reductive transamination to the corresponding amino acid. In addition to the unmodified amino acids, I also found numerous N methyl and formyl modifications, as well as amino acid oligomers and oligomers containing N-modified peptides (Figure 79). These species were identified by exact molecular weight to within 10ppm and by comparing fragmentation patterns to literature values as appropriate.

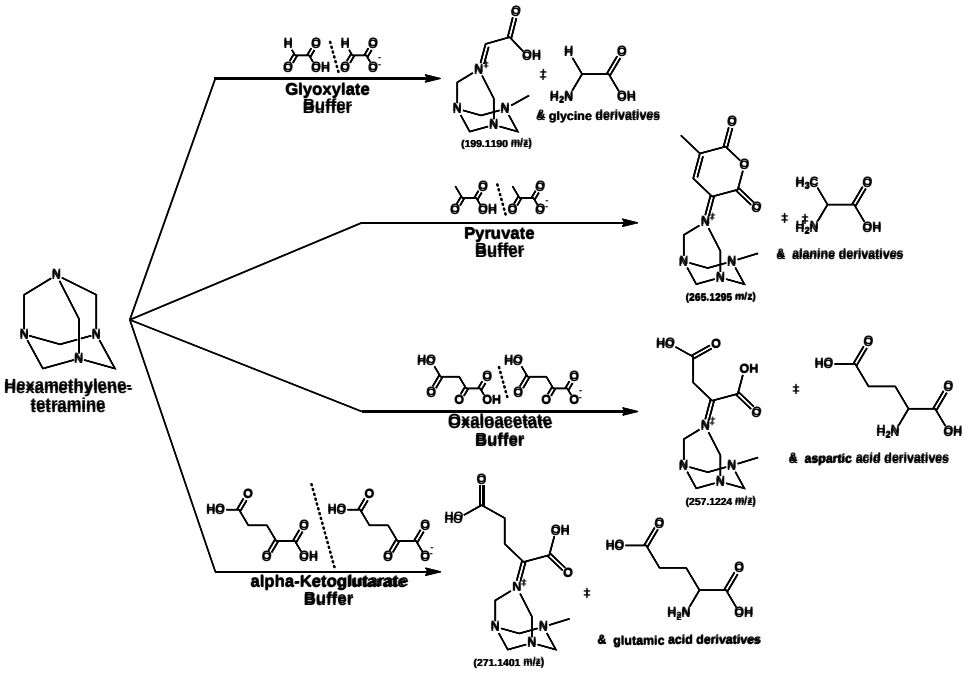


Figure 79 – Products of Reaction of HMTA with α-Keto Acids. Reaction of HMTA and α-keto acids resulted in the formation of HMTA-acid adducts and subsequent reduction to free or N-modifed amino acids, as well as peptide oligomers.

This work is not the first time HMTA has been investigated in terms of peptide formation. Fox and Windsor investigated the reaction of formaldehyde and ammonia under reflux conditions and observed the production of HMTA along with seven amino acids[2]. Further experimentation by Wolman and Miller showed small yields of amino acids after holding an aqueous solution of HMTA at reflux and subsequent acid

hydrolysis[3]. Neither Fox and Windsor, nor Wolman and Miller reported any peptide oligomer formation.

Based on these observations, I proposed that, in an aqueous solution high in peptide concentration, HMTA can act as a coupling agent for peptide bond formation in a catalytic cycle (Figure 1). To test this hypothesis, we heated an aqueous solution of 0.5 M glycine solution with 0, 0.1, and 1.0 equivalents of HMTA to glycine to 100°C for twelve days and analyzed the products using liquid chromatography mass spectrometry (LCMS). Our initial analysis entailed qualitative confirmation of m/z values corresponding to glycine oligomers as well as the corresponding HMTA ester adduct **3** formed from the reaction of glycine (or glycine oligomers) with the methyl-HMTA adduct **2**. Qualitative LCMS data confirmed the presence of these compounds, and shows an increase in signal of each investigated oligomer and adduct with increased equivalents of HMTA. Herein, we present the findings of our investigation into the role of HMTA in a catalytic cycle for the formation of glycine oligomers.

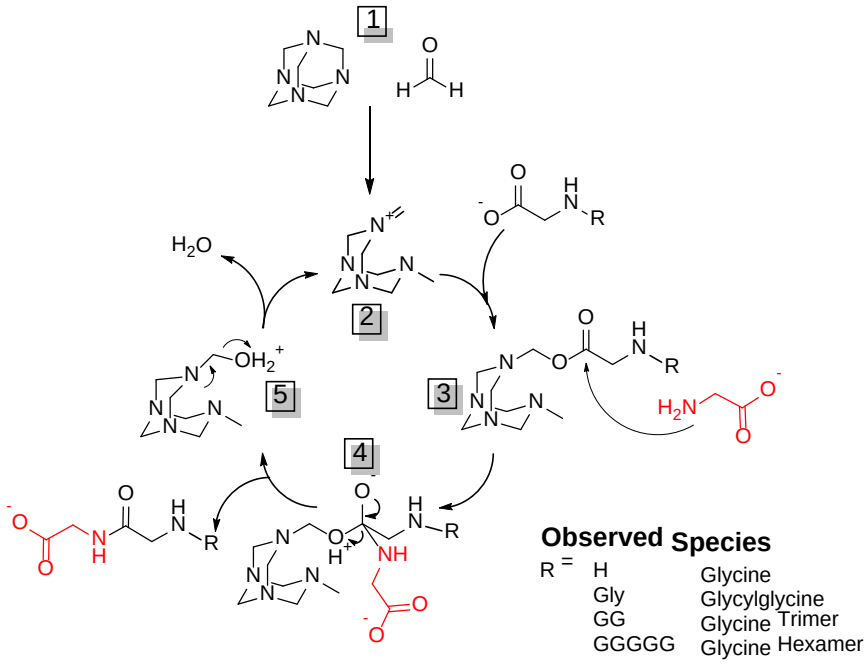


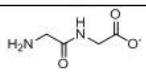
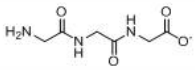
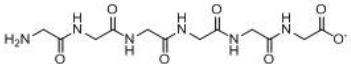
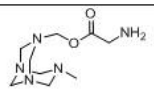
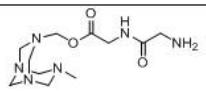
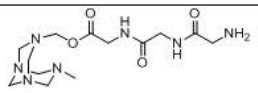
Figure 80 – Proposed Catalytic Cycle for the Formation of Glycine Oligomers via Hexamethylenetetramine. HMTA and free formaldehyde **1** in solution react to form a positively charged iminium **2**. The carboxylate moiety can undergo a nucleophilic addition to the methylene group of **2**, yielding a neutral ester adduct **3**. We propose that formation of this ester activates the carboxy group of glycine by creating a leaving group with superior leaving potential to hydroxide or water under these conditions. **3** is attacked at the carbonyl moiety by the lone electron pair on the amine nitrogen of a second glycine or glycine oligomer forming an unstable, alkoxide **4**. This alkoxide intermediate fragments to a hydroxymethyl-HMTA adduct **5** and the observed glycine oligomer. Dehydration of **5** yields **2** to restart and continue the catalytic cycle.

Results

Species shown in Table 7 were chosen for analysis solely based on commercial availability of pure standards for the given glycine oligomer. The compounds chosen were glycine, glycine diketopiperazine (DKP), glycylglycine (GG), glycylglycylglycine (GGG), and glycylglycylglycylglycylglycylglycylglycine (GGGGGG). We believe that the trends observed in these oligomers are representative of all glycine oligomers. Expected reaction products and their corresponding m/z values are shown in Table 1. Qualitative LCMS results were obtained and are illustrated in Figure 81. There is evidence of all expected products and a trend of increasing relative product concentration with increasing HMTA molar ratios is observed, with the exception of glycine. The

decrease in relative concentration of glycine is to be expected due to the consumption of the monomer to yield oligomers in the proposed catalytic cycle. Method development to obtain quantitative data is in progress.

Table 7 – Expected Products in the Proposed Catalytic Cycle for the Formation of Glycine Oligomers Via HMTA.

Expected Product		m/z
Glycine (Gly; G)		76.0394
Glycine DKP		115.0503
Gly-Gly (GG)		133.0608
Gly-Gly-Gly (GGG)		190.0823
Gly-Gly-Gly-Gly-Gly-Gly (GGGGGG)		361.1467
HMTA + Formaldehyde		155.1292
HMTA-Gly Adduct		230.1612
HMTA-GG Adduct		287.1827
HMTA-GGG Adduct		344.2041

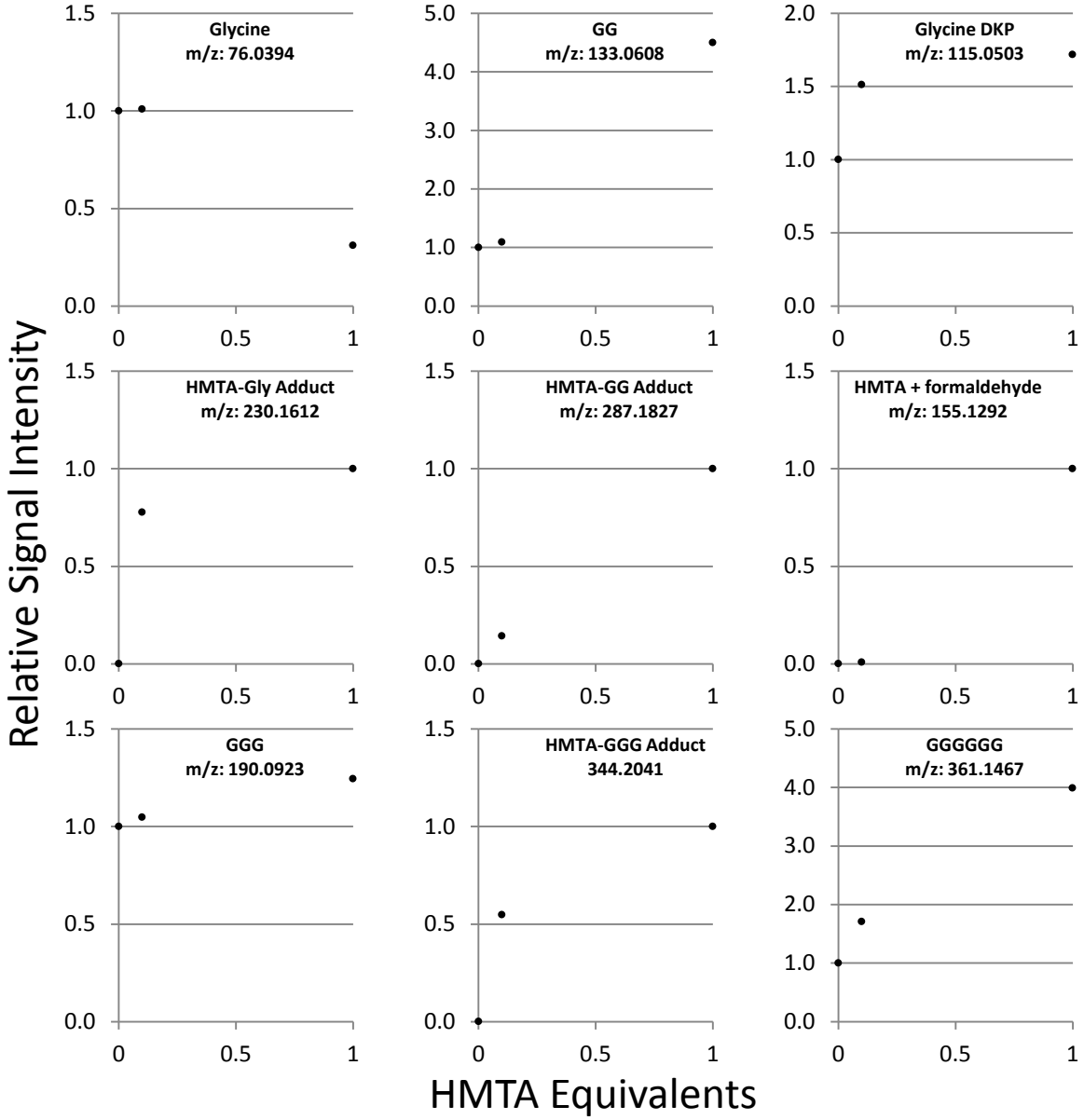


Figure 81 – Relative Signal Intensity of Glycine and Glycine Oligomers as well as Selected HMTA Adduct Species. Each reaction above contained 0.5M glycine and 0, 0.1, or 1 equivalent of HMTA at a pH of 5.9. Solutions were heated for 12 days and analyzed by LCMS resulting in the data shown above. Comparison of the plots shows that with increased HMTA concentration glycine is depleted, and all glycine oligomer concentrations are increased. This is further substantiated by increased signals of all HMTA-Glycine adducts with increasing concentration.

In our analysis, we assume equal ionization efficiency in each medium. Our findings indicate that oligomers are present in increased concentration in the presence of HMTA; consistent with the proposed catalytic cycle and further supported by the presence of signals consistent with HMTA-peptide oligomer adduct species. We

anticipated a large presence of glycine DKP due to the kinetic favorability of cyclization of the glycine dimer[4]. Surprisingly, the increase in signal due to DKP is less than the increase in linear dimer, suggesting that formation of the HMTA ester adduct is rapid and outcompetes the DKP ring closure. These results are the first evidence of a catalytic role of HMTA in peptide oligomer formation.

Recommendations: Quantitative yields for products of the reaction of HMTA with glycine are necessary to support the qualitative findings of this project. Quantification of the glycine oligomers will be achieved by implementing multiple reaction monitoring mass spectrometry (MRM-MS), to provide quantitative data.

A.1 Experimental

Preparation: All reagents utilized were obtained from Sigma Aldrich at the highest available grade unless otherwise noted. A 0.5 M stock solution of glycine was prepared using 100-mL of LCMS-grade water. Three reactions each at a 1:1 and 10:1 molar ratio of glycine to HMTA were prepared using 10 mL of the glycine stock solution per reaction. Control reactions consisted of three 10-mL samples of the glycine stock solution and three samples of 0.5 M HMTA in 10 mL of LCMS-grade water.

The pH of the glycine stock solution was measured to be 5.90 using a glass pH meter. After addition of HMTA, the pH was adjusted drop wise to 5.90 ± 0.04 using a 0.1 M hydrochloric acid (HCl) stock solution.

Conditions: The reactions were refluxed at 100°C with stirring for twelve days. A carousel reactor was used to run all twelve experiments in conjunction under the exact

same conditions. After twelve days, the reactions were transferred from the reaction vessels to vials and stored at 10°C for future LCMS sample preparation.

Analytical Methods: LCMS samples of the reactions were prepared at a 100:1 dilution using LCMS-grade water. Analysis was conducted using a Waters 3100 SQD MS set to positive electrospray (M+1) with direct sample injection. Data was acquired through the instrument's accompanying MassLynx software and interpreted using Microsoft Office Excel. Species were confirmed by referencing appropriate mass-to-charge (m/z) ratios for the positive ions of expected products.

A.2 References

[1] WALKER J. Formaldehyde. Formaldehyde. 3rd ed. New York: Reinhold; 1964.

[2] FOX SW, WINDSOR CR. Synthesis of Amino Acids by the Heating of Formaldehyde and Ammonia. Science 1970;170:984-6.

[3] WOLMAN Y, et al. Formaldehyde and Ammonia as Precursors to Prebiotic Amino Acids. Science 1971;174:1038-41.

[4] SAKATA K, et al. Effects of pH and temperature on dimerization rate of glycine: Evaluation of favorable environmental conditions for chemical evolution of life. Geochimica et Cosmochimica Acta 2010;74:6841-51.

APPENDIX B
CONVERSION OF FRUCTOSE TO 5-HYDROXYMETHYLFURFURALDEHYDE IN BUTADIENE SULFONE

The DOE "Biomass" publication "Top Value Added Chemicals from Biomass," [1] identifies 2,5-Furandicarboxylic acid(FDCA) and levulinic acid(LA) as two potential building blocks for a biomass based refinery; 2,5-Furandicarboxylic acid could replace terephthalic acid in polymerization reactions as well as be converted to other useful molecules(Figure 82) while levulinic acid can act as feed stock to produce methyl tetrahydrofuran, and other molecules (Figure 83). Both FDCA[2] and LA[3] can be produced from 5-hydroxymethal 2-furfuraldehyde(HMF), (Figure 84,Figure 85)

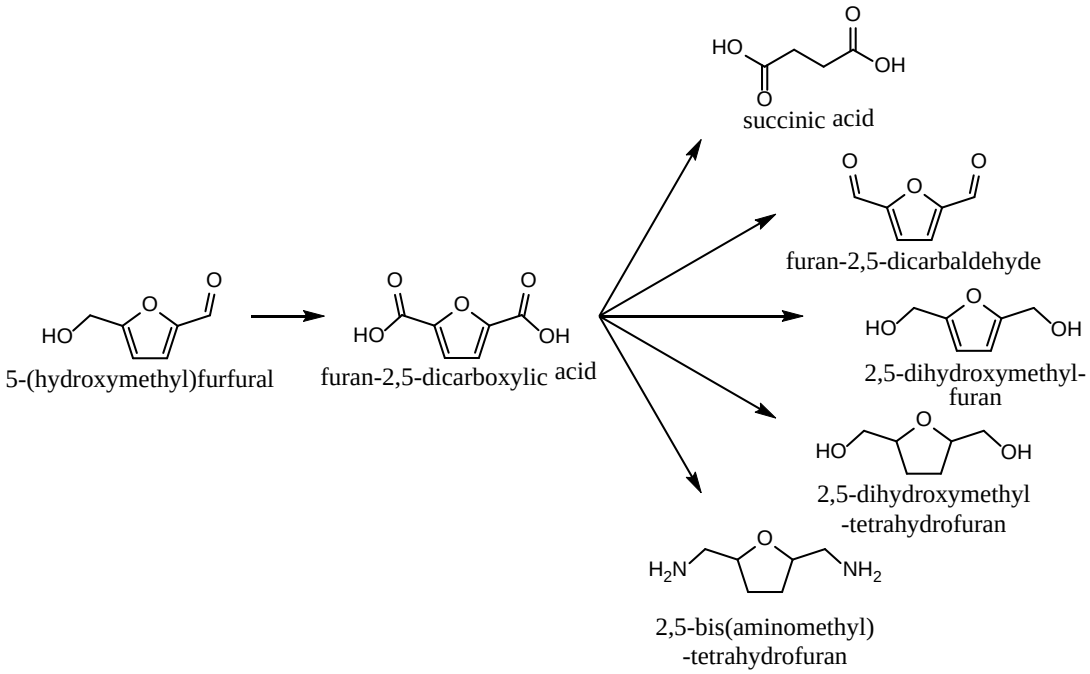


Figure 82 - Reactions of HMF to other molecules. FDCA could be bypassed for many of these molecules. (Adapted from Werpy)

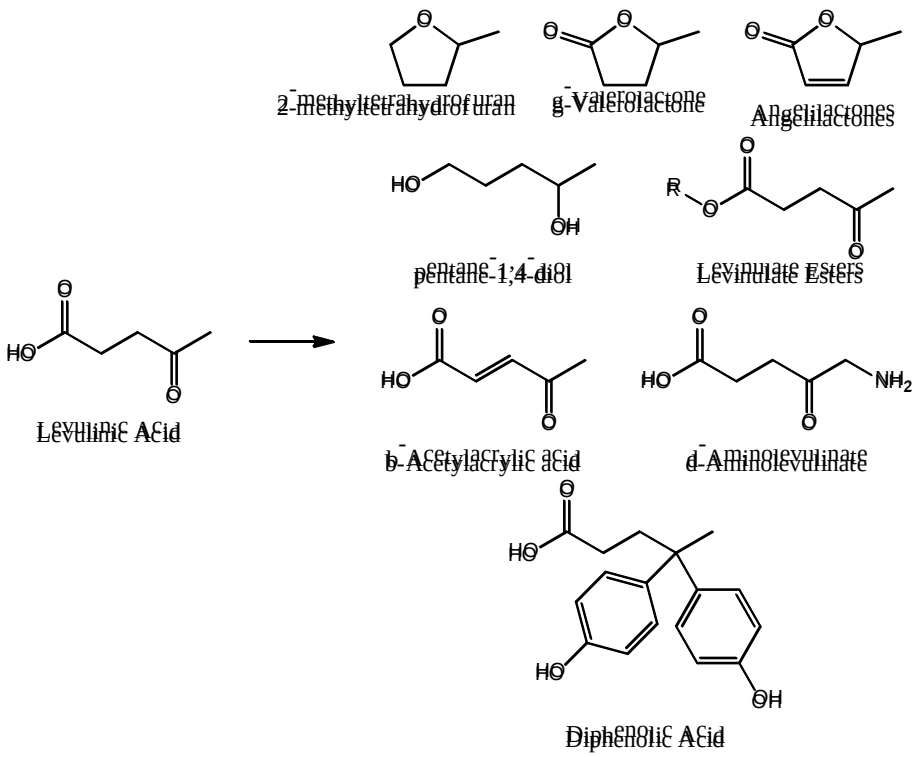


Figure 83 - Potential conversions of Levulinic Acid based on known chemistry. (Adapted from Werpy)

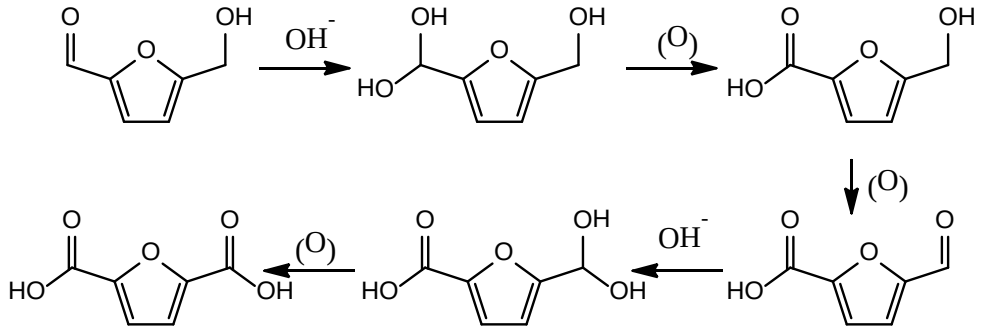


Figure 84 - Oxidation of HMF to FDCA. Adapted from Casanova.

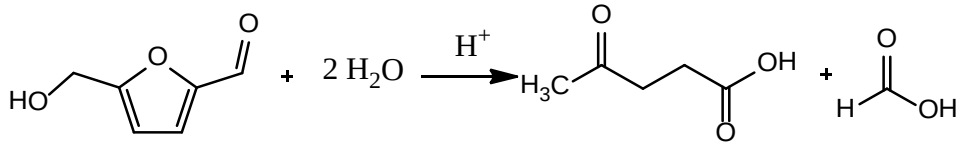


Figure 85 - Conversion of HMF to LA. Adapted from Girisuta.

HMF can be produced from fructose in DMSO in high yields, however the separation is difficult, and wasteful. For example, Roman-Leshkov and Chheda show production of HMF from fructose, glucose, and inulin, with good yields (>70% for fructose and inulin~50% for glucose), using a system of DMSO/H2O with a methylisobutylketone/2-butanol phase present in the reactor for product extraction[4, 5]. Tuerck, et al. showed ~85% yield in the same solvent system using a microreactor system[6]. Halliday showed 85% HMF yield in pure DMSO using a Dowex ion exchange resin[7]. However in each of these cases, despite the good yields, recovery of the product requires expensive liquid liquid extraction processes, rendering the reaction impractical.

To address the difficulties associated with separating HMF from DMSO, we proposed to use piperylene sulfone, a solvent with very similar solvatochromic properties to DMSO (Figure 86). These similarities make piperylene sulfone a viable replacement as reaction medium for many reactions which are carried out in DMSO. Piperylene sulfone is formed through a reverse cheletropic reaction between sulfur dioxide and 1,3-trans-piperylene. Superiorly to DMSO, PS can be decomposed to SO2 and piperylene at temperatures of ~110C (Figure 87) which allows for facile separation of reaction products (Figure 88)[8]. The off-gassed sulfur dioxide and trans-piperylene can be re-reacted elsewhere to reform the piperylene sulfone, enabling recyclability.

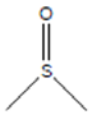
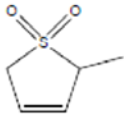
Property	*DMSO* 	*Piperylene Sulfone* 
Boiling Point (°C)	189	decomposes
Melting Point (°C)	16-19	-12
α	0	0 ± 0.1
β	0.76	0.46 ± 0.08
π^*	1.00	0.87 ± 0.04
$E_t(30)$ (kJ/mol)	189	189 ± 0.3
ε	46.7	42.6

Figure 86 - Comparison of solvatochromic properties of piperylene sulfone and DMSO.

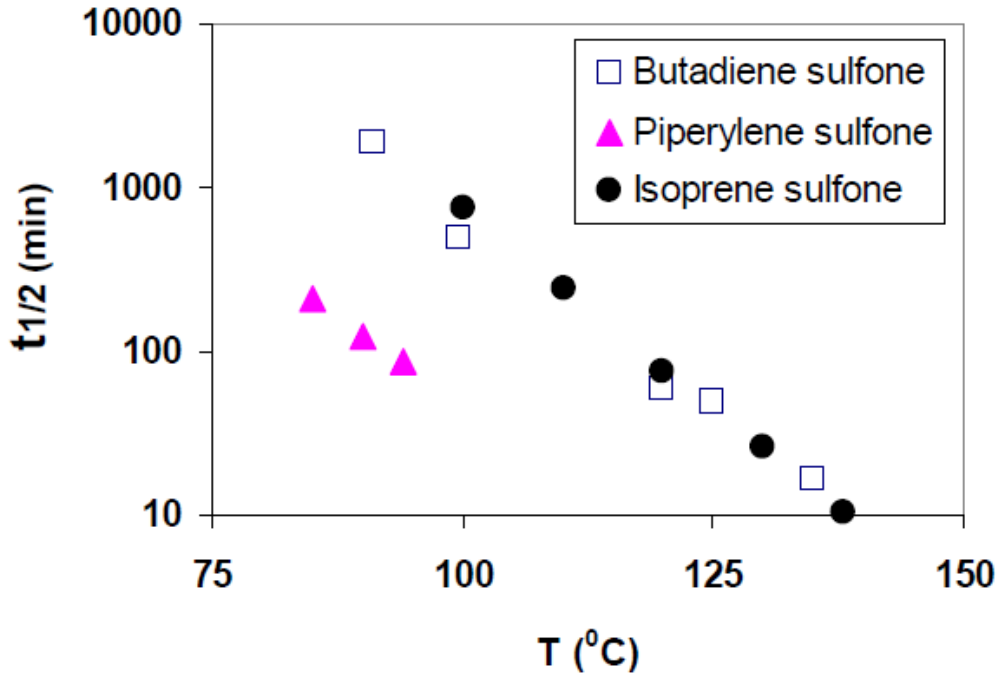


Figure 87 - Decomposition Half-Lives of various sulfones as a function of temperature.

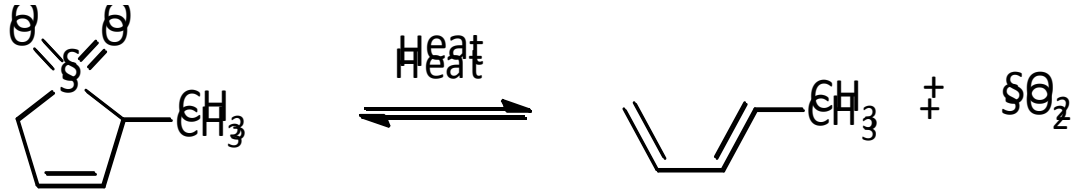


Figure 88 - Decomposition of piperylene sulfone.

Our lab has previously investigated the reaction of fructose to form HMF in piperylene and butadiene sulfones. We have used butadiene sulfone as a commercially available analog to piperylene sulfone. The main difference being the decomposition temperature of butadiene sulfone is higher. (Figure 87). Initial reactions of HMF in butadiene sulfone were conducted by heating .1M fructose at 110^{o}C in a 4:1 mixture of butadiene sulfone and water, for times ranging from 2 to 24 hours. It was believed that decomposition of the butadiene sulfone and subsequent reaction of the liberated sulfur dioxide would yield sulfonic acid; the sulfonic acid would then catalyze the dehydration of fructose to HMF via the traditional acid catalyzed dehydration pathway. However, these reactions had only modest yields of 10-20% [8].

Our lab has also shown that an HMF standard can be separated from piperylene sulfone by heating and vacuum with quantitative yield. Separation was conducted at 95 ^{o}C in a Buchi Kugelrohr under an inert flow and in the presence of up to 20% water, without negative effect. Separation was shown to be pure by both mass balance and NMR. No products of degradation were observed in the NMR [8].

As a potential bulk chemical a 20% yield was far too low to be considered for further improvement. Some other catalyst was needed to improve the reaction. As the sulfonic acid showed some viability as a catalyst, we turned our attention toward catalysts containing sulfonic acid groups. Shimizu, et al. [9] have shown 100% yields of HMF from up to 50 weight percent fructose in DMSO using Amberlyst 15 as a catalyst.

Amberlyst 15 is a solid acid resin consisting of sulfonic acid groups bound to a highly cross-linked styrene-divinylbenzene copolymer[10]. Despite the high production possible from this reaction, dependence upon DMSO complicates the separation stage. Combination of PS as solvent and Amberlyst-15 as catalyst could allow high yield reaction with trivial separation.

To evaluate optimal conditions for HMF production we performed a statistical design of experiments based on the method described by Louvar.[11] We chose temperature, time, catalyst loading and particle size, sugar concentration, and rate of stirring as the variables to analyze in our design of experiments. The preliminary results from our design of experiments showed an 85% yield of HMF at high temperature (120^{o}C), and 70% yield at the lower temperature (100^{o}C).

Based on the results of the design of experiments and the limitations of piperylene sulfone at temperatures upwards of 100 degrees, we chose the following conditions for our optimized run: 100^{o}C, 4mg/mL of catalyst loading, .125mm catalyst diameter, 45 minutes, 100 rpm stirring, and .1 molar fructose concentration. Unfortunately, when running these reaction in piperylene sulfone, humin formation became the prevalent reaction, greatly decreasing the reaction yields. We were unable to reduce this sideproduct, despite the fact that ^{13}C NMR of the supernatant showed only PS and HMF.

Table 8 – Yields of Desigin of Experiments Reactions in Butadiene Sulfone

Run #	Temp. (°C)	Time (min)	Catalyst loading (mg/mL)	Catalyst Diameter (mm)	Sugar Conc. (M)	Mixing Rate (rpm)	Yield (%)
1	100	30	2	0.063	0.1	120	34.76
2	120	30	2	0.063	0.5	120	17.96
3	100	60	2	0.063	0.5	180	15.65
4	120	60	2	0.063	0.1	180	75.71
5	100	30	4	0.063	0.1	180	74.47
6	120	30	4	0.063	0.5	120	19.30
7	100	60	4	0.063	0.5	120	24.29
8	120	60	4	0.063	0.1	180	85.74
9	100	30	2	0.125	0.5	180	16.72
10	120	30	2	0.125	0.1	120	80.37
11	100	60	2	0.125	0.1	120	86.76
12	120	60	2	0.125	0.5	120	35.68
13	100	30	4	0.125	0.5	120	26.58
14	120	30	4	0.125	0.1	120	88.22
15	100	60	4	0.125	0.1	180	23.57
16	120	60	4	0.125	0.5	180	31.60
17	110	45	3	0.125	0.5	150	15.29
18	110	45	3	0.125	0.3	180	22.02
19	110	60	3	0.094	0.5	150	18.61
20	120	45	3	0.094	0.5	150	29.47
21	110	45	4	0.094	0.3	180	45.99
22	110	60	3	0.094	0.3	180	38.55
23	110	45	3	0.094	0.5	180	22.57
24	120	45	3	0.094	0.3	180	40.56
25	110	45	3	0.094	0.3	150	46.32

While we were unable to optimize the reaction in piperylene sulfone, the result presented here illustrates the potential for acid catalyzed reactions conducted in DMSO to be converted to butadiene sulfone. While separation of HMF from butadiene sulfone is impractical due to the decomposition to levulinic acid, use of BS rather than DMSO may be an energy saving proposition in a system with a more thermally stable product.

B.1 Experimental

Reactions of Butadiene Sulfone: Solid butadiene sulfone was weighed into a 25mL round bottom flask. To this, the necessary mass of catalyst was added followed by the necessary sugar mass and a stir bar. A septum was added to the flask, and nitrogen added by flow through two needles punctured through the septum. The flask was then immersed to the neck in an oil bath at the desired temperature.

Catalyst sizing: Catalyst of the desired size was created by running stock Amberlyst through a burr-type coffee grinder. The ground catalyst was then run through a Sigma Aldrich Mini-sieve micro sieve set using 80, 120, 170, and 230 mesh screens to obtain catalyst with the desired size distribution.

Analysis: Analysis was conducted using a Waters 3100 Mass Detector attached to a 2696 separations module. LC separation was conducted on a C18 column with a water/acetonitrile gradient method.

B.2 References

[1] AUTHOR. Top Value Added Chemicals from Biomass: Volume I -- Results of Screening for Potential Candidates from Sugars and Synthesis Gas. Other Information: PBD: 1 Aug 20042004. p. Medium: ED; Size: 76 pp. pages.

[2] CASANOVA O, et al. Biomass into Chemicals: Aerobic Oxidation of 5-Hydroxymethyl-2-furfural into 2,5-Furandicarboxylic Acid with Gold Nanoparticle Catalysts. ChemSusChem 2009;2:1138-44.

[3] GIRISUTA B, et al. A kinetic study on the decomposition of 5-hydroxymethylfurfural into levulinic acid. Green Chemistry 2006;8.

[4] ROMÁN-LESHKOV Y, et al. Phase Modifiers Promote Efficient Production of Hydroxymethylfurfural from Fructose. Science 2006;312:1933-7.

[5] CHHEDA JN, et al. Production of 5-hydroxymethylfurfural and furfural by dehydration of biomass-derived mono- and poly-saccharides. Green Chemistry 2007;9:342-50.

[6] TUERCKE T, et al. Microreactor Process for the Optimized Synthesis of 5-Hydroxymethylfurfural: A Promising Building Block Obtained by Catalytic Dehydration of Fructose. Chemical Engineering & Technology 2009;32:1815-22.

[7] HALLIDAY GA, et al. One-Pot, Two-Step, Practical Catalytic Synthesis of 2,5-Diformylfuran from Fructose†. Organic Letters 2003;5:2003-5.

[8] MARUS G. Application of Green Chemistry and Engineering to Novel Sustainable Solvents and Processes: Georgia Tech; 2011.

[9] SHIMIZU K-I, et al. Enhanced production of hydroxymethylfurfural from fructose with solid acid catalysts by simple water removal methods. Catalysis Communications 2009;10:1849-53.

[10] KUNIN R, et al. Characterization of Amberlyst 15. Macroreticular Sulfonic Acid Cation Exchange Resin. I&EC Product Research and Development 1962;1:140-4.

[11] AUTHOR. Simplify Experimental Design. Chemical Engineering Progress. New York, NY, ETATS-UNIS: American Institute of Chemical Engineers; 2010. p. 6.

VITA

Butch grew up in Horicon, Wisconsin and attended the University of South Carolina, on a McNair Scholarship, and met his wife Diane. He graduated with a Bachelors of Science in Chemical Engineering. After undergraduate, he worked for two years for the Defense Nuclear Facilities Safety Board where he gained experience with chemistry of nuclear materials separations, and radiolytic chemistry. When not in the laboratory, he enjoys playing board games, and explaining scientific concepts using cat analogies on Reddit.